TITLE

Gender Wage Differentials and the Labour Market for Young Workers:

An Empirical Analysis Using Data for Ireland

AUTHOR

Barry Reilly

INSTITUTION and DATE

University of Warwick

March 1989

Gender Wage Differentials and the Labour Market for Young Workers:

An Empirical Analysis Using Data for Ireland

Barry Reilly

Thesis Submitted for Ph.D

Dept. of Economics

University of Warwick

March 1989

Table of Contents.

Table of Contents (Cont'd)

Table of Contents (Cont'd)

List of Table Titles.

List of Table Titles (Cont'd).

Acknowledgments

I am gratefully indebted to both my supervisors Ben Knight and Mark Stewart for their supervision and guidance without which this work would never have been completed. This thesis has also benefited greatly from the comments and suggestions of Wiji Narendranathan and Brian Chiplin. Useful comments have also been provided on different chapters by Martyn Andrews, Anne Gibbons, Claudio Lucifora, Steve Machin, Robin Naylor, Brendan Walsh, Chris Williams and participants of the Labour Economics Workshop at the University of Warwick. The assistance provided by Gerry Sexton, Brendan Whelan and James Williams of the ESRI in providing the data set used in the analysis is also gratefully acknowledged as is the computational assistance provided at different times by Keith Halstead, Peter Mitchell and David Ross at the University of Warwick. The funding provided by an Economic and Social Research Institute Fellowship between October 1985 and September 1987 and by a University of Warwick Department of Economics Tutorial Fellowship from October 1987 to September 1988 is also gratefully acknowledged. Needless to say, in the absence of either of these Fellowships completion would have been difficult.

Summary

Little empirical evidence is available on the determinants of wages at the level of the individual for Ireland and to the author's knowledge no such evidence is available for young workers. One objective of this thesis, therefore, is an examination of the determinants of wages at the micro level using data from a national survey of young workers recently undertaken in Ireland. The effects of education, training, region, industry and occupation are assessed. More importantly, gender wage differentials are calculated for the sample under a number of alternative assumptions concerning the treatment of occupations. In the applied econometric literature relating to the estimation of both wage equations and gender wage differentials little emphasis has focused on the appropriate treatment of occupations. In view of this, an econometric objective of this thesis is an analysis of how the gender wage differential is affected by altering the econometric assumptions underlying occupations. The sensitivity of the gender wage differential to occupational endogeneity is examined in a dichotomous framework using two contrasting econometric methods. Statistical tests for occupational exogeneity are provided and their results reported. Structural occupational models are also estimated. To assess the effects of occupational segregation on the gender wage differential a five-way occupational categorisation is employed and an effort is made to disentangle inter and intra occupational wage effects. Occupations are agained treated as endogenous and a consistent estimator designed to correct for selectivity bias is employed. In both the dichotomous and the polychotomous frameworks the estimated gender differentials appear sensitive to occupational endogeneity. Finally, the issue of segregation is again addressed but this time in the context of the dual labour market. An empirical dual labour market model is estimated using an endogenous switching model with partial observability in the latent dependent variable. Sectoral differentials are calculated and the results of an informal test of rationing, a basic tenet of dual labour market theory, tentatively suggest that primary sector rationing, to the extent it exists, falls disproportionately on the young females in the sample.

Chapter One

Introduction

This thesis is motivated by economic and econometric considerations. The economic objective is the analysis of wage determinants at the level of the individual in the labour market for young workers in Ireland. Particular emphasis is placed on quantifying the residual gender wage gap that may be determined by, among other factors, the exercise of discriminatory power. The existence of discrimination has clear policy implications and implies that female wages are conditioned unfairly on the basis of their gender. The particular focus on young workers is also of some interest since it could be argued that both males and females have had relatively homogeneous and uninterrupted labour force experiences.

To the author's knowledge the above has been the subject of no attention to date in the Irish context. Even for the adult labour market in Ireland little research effort has been expended on analysing the determinants of adult earnings[1]. An exception in this regard is Walsh and Whelan (1976) who used a sample of redundant adult workers to estimate human capital earnings functions in an attempt at analysing gender differentials in earnings. As pointed out above the emphasis in this thesis is on young workers using a data set derived from a relatively recent national survey (1982). A full description of the data set used is reserved until chapter three but suffice for it to state here that is more general than that of Walsh and Whelan (1976) and could be interpreted as containing fewer of the biases associated with their data.

The absence of a body of research is a clear motivation in itself. However, it should not be interpreted as the sole justification for embarking on such an exercise. Though the results reported in chapters four to seven possess a clear Irish relevance the econometric methodology em-

[1] This is in large part explained by either the unavailability or lack of access to suitable data sets.

ployed has a much broader relevance. Though this is to be addressed more extensively in chapter two the standard approach[2] to the estimation of the discrimination effect involves the estimation of earnings or wage equations by OLS. The discrimination estimate is obtained as the residual difference in observed wages that persists having controlled for a standard set of productivity and compensating variables. The earnings or wage equations are invariably reduced form and the discrimination estimates are interpreted as unconditional. Three broad econometric problems are invariably ignored when such an approach is adopted. The first is to some extent dictated by data considerations and the lack of adequate female labour force information. Potential rather than actual labour force experience is used for females and this has been shown to lead to an overstatement of the discrimination estimate (see chapter two).

The second problem concerns the fact that the observed sample of females may not be a random sample of females drawn from the population as a whole. This truncation effect has clear implications for the wage equation estimates and hence the discrimination estimates. Some authors have attempted to deal with this problem by examining only single male and female individuals. However, these may not be the appropriate comparator group. Furthermore, focusing exclusively on this group provides a relatively narrow definition of gender wage discrimination. Consistent estimators (see chapter two) have been used in a number of studies to account for the selectivity bias associated with participation.

The third problem is the treatment of occupations. Failure to explicitly account for it in wage equations allocates the compensating differential effects to the residual. This leads to a higher discrimination effect than may actually exist. Allowing occupational levels to enter exogenously may lead to biased wage coefficient estimates if, in fact, occupations are endogenous.

In the context of the data set used in this thesis the first two issues are not interpreted as being problematic. Since the survey questionnaire allows the precise calculation of labour force experience (see chapter three) no artificial constructs are required for this key variable. The selectivity bias effects that may come through participation is not deemed an important issue since the

[2] This approach owes a lot to Oaxaca (1973) and is termed the "index number" approach.

vast majority of the young workers (male and female) surveyed are unmarried. Furthermore, all those surveyed described themselves as either working or actively searching for work and though no results are presented here no statistical evidence to suggest that the sample of workers are systematically different from the population of workers and non-workers as a whole was found[3].

The issue of potential occupational endogeneity provides the second and perhaps more important econometric motivation for the thesis. Little research has been directed at analysing the effects of occupational endogeneity on the discrimination estimates. The inadequate treatment of occupations as exogenous intercept shifts in wage equations fails to account for a channel through which discrimination may in fact operate. If females are sample selected by employers or self-select themselves into certain occupations then the treatment of occupations as exogenous is clearly invalid [4]. Thus a major thrust of this study is an investigation of how occupational endogeneity affects discrimination estimates with a particular application to the labour market for young workers in Ireland. Statistical tests for occupational exogeneity are also provided.

How this study proposes to achieve this twin set of objectives may best be elucidated by briefly outlining the layout of the remainder of the thesis. The following chapter, chapter two, provides a separate theoretical and empirical survey of the relevant discrimination literature. Though for the most part economic theoretical considerations play a peripheral role in the subsequent analysis the importance of a theoretical framework inside which to analyse the results should not be under-stated. Nevertheless, without pre-judging matters it is clear that the phenomenon of discrimination (like many other economic phenomenon) fails to provide a dominant theoretical paradigm that can be conveniently invoked. Therefore, of more importance

---

[3] There are a number of channels through which selectivity bias could potentially operate. The observed sample of workers could be a non-random sample as a consequence of participation selectivity bias. Futhermore, given the terminal age of twenty-four for inclusion in the sample those still in higher education are excluded thus providing another potential source of selectivity bias. To simultaneously model these different forms of selectivity bias would be an immense and complex task. It is for this reason that these forms of selectivity bias are conveniently ignored allowing attention to focus exclusively on the effects of occupational endogeneity. Though no evidence is presented the effects of participation and education selectivity bias are thus assumed negligible.

[4] Rational expectations may be invoked to explain the behaviour of both the self-selecting females and the sample-selecting employers. If females anticipate absence from the labour force they may optimally choose occupations with relatively low-skill levels. On the other hand, employers may optimally choose to invest in those workers whose attachment to the labour force is more permanent in character. This will be particularly so in terms of firm specific training paid for by the firm.

from the point of view of the analysis presented is the survey of the empirical evidence which also surveys the relevant econometric issues and places the issues to be presently addressed in context.

Chapter three provides an extensive description of the sample survey and data set used in the analysis pointing out some of its advantages and its potential limitations. The Irish labour market context inside which the survey took place is also briefly sketched.

Chapter four[5] presents the first analysis of the data. In this chapter the traditional "index number" methodology is outlined along with its limitations. Standard reduced form wage equations are estimated allowing inferences to be made regarding the returns to educational and vocational qualifications and labour force experience *etc*. Occupations are treated as exogenous intercept shifts in this chapter. Gender wage differential estimates (with associated standard errors) conditional and unconditional on occupations and industries are reported. Mean and base gender wage differential estimates and deviations from the base are also reported.

Chapter five explicitly treats occupational attachment as endogenous where the occupational categorisation is on the basis of a manual/non-manual dichotomisation. Two consistent estimators (the Heckman two-step estimator) and a distribution-free estimator based on instrumental variables (IV) are contrasted with the standard OLS estimator. Tests of occupational exogeneity are provided within both the Heckman and the IV framework. Unexplained gender wage differentials are reported for each of the two occupational categories. The estimates are shown to be sensitive to both occupational endogeneity and the manner in which the endogeneity is accounted for. Results based on the estimation of structural occupational attachment models are also reported.

The gender wage differential and the influence exerted on it by the occupational segregation of females into low-paying occupations has been recently examined in the literature (see chapter two). In both the US and the UK evidence has been presented to suggest that the female wage

---

[5] A version of chapter four has already been published in the Economic and Social Review (see bibliography).

disadvantage is explained more by their intra-occupational wage disadvantage than by their segregation into certain occupational categories. Chapter six examines this particular issue in the context of young workers in Ireland. The econometric innovation that distinguishes this chapter from previous studies of its kind is that the occupational wage equations are corrected for selectivity bias. The findings, however, are broadly in line with what has been discovered in both the US and UK.

The final empirical chapter, seven, departs to some extent from the framework used in chapters four to six and focuses on the dual labour market phenomenon. However, a linking theme still remains since the issue of gender is very much to the fore. One objective of this chapter is to establish the extent to which gender is used as a mechanism to allocate workers between primary and secondary sectors. Occupations are also used to allocate some individuals to primary and secondary sectors. A stochastic switching model with partial observability in the latent dependent variable is used to estimate the parameters of the dual labour market model and wage gaps between the sectors are estimated with special reference to gender. A tentative test of primary sector rationing is also proposed.

Finally, chapter eight attempts to draw the diverse strands of chapters four to seven together in an attempt at presenting conclusions.

Chapter Two

Discrimination and the Gender Wage Gap: A Theoretical and Empirical Survey

2.1 Introduction

The existence of a gender based wage differential has exercised the interest of theoretical and empirical economists alike over the past number of decades. Even at the start of this particular decade a sizable gender wage differential still existed in the twelve industrialised countries surveyed by Mincer (1985). Though the wage gap has closed it still remains a strong statement to say that most of the remaining differential is linked to gender differences in life-time work experience. Invariably two competing explanations are invoked to explain the persistence of the gender wage gap. The human capital explanation places emphasis on differentials in the accumulation of human capital investments which leaves women less productive than men and hence through marginal productivity conditions with lower wages than men. The second explanation focuses on the payment of differential wages to workers of comparable productivity where the differential is determined by gender.

This survey is divided between a theoretical and an empirical survey of the relevant literature. The theoretical models that have been suggested in the literature can broadly be assigned to either supply side or demand side explanations. The theoretical section sub-divides into a brief survey of the supply side explanations and a more comprehensive survey of the demand side explanations. The demand side section is itself divided into a number of sub-sections that focus separately on neoclassical, institutional and Marxian paradigms[6]. It is within this section that both the economics of discrimination and the theory of dual labour markets are explored. The

[6] The allocation of the Marxian paradigm to this particular category is more a convenience than a statement.

empirical survey section concentrates more extensively on the direct and indirect tests of discrimination and the econometric problems of estimating both the gender wage gap and testing for the existence of dual labour markets.

The Gender Wage Gap: A Theoretical Survey

2.2 Supply Side Explanations

The supply side explanations have their origin in the human capital theory as developed by Becker (1975) and Mincer (1974). Using a human capital explanation Mincer and Polachek (1974) attempt to explain observed gender differentials in life-cycle earnings in terms of gender differentials in the volume of accumulated human capital investments. The labour force intermittency of females dictated by the timing profile of child birth *etc.* has implications for the incentives and opportunities available to females in acquiring productivity enhancing human capital investments.

The above explanation has been criticised, however, for its failure to treat human capital investments as heterogeneous. Polachek (1981) outlines a theoretical model that allows for different types of human capital investments. The "hedonic price" approach outlined in Rosen (1974) is used to embed occupational choice into the human capital framework. Thus, human capital theory is used to explain the determinants of occupational structure. Crucial to the Polachek (1981) account is the role played by the atrophy rate or depreciation rate associated with different skill levels. In the utility maximising framework, females are assumed to optimally choose both the volume and type of human capital in order to maximise life-cycle earnings. Females are assumed, in anticipation of labour force interruption, to choose those occupations that are characterised by low atrophy or depreciation rates. This may eventuate in their segregation into low skill and low wage occupational categories.

More intimately linked to both the supply side and to human capital theory described above are the explanations described in terms of the allocation of time. Originally introduced in response to deficiencies identified in the traditional theory of labour supply (see Killingsworth (1983)) they have been used to explain gender differences in labour supply and hence gender differences in earnings and wages. The seminal contribution in this area is again due to Becker (1965) and central to the analysis is the concept of the household production function. The formulation is in line with the characteristics approach of Lancaster (1966) where time and market goods are treated as inputs in household production. Gronau (1977) provides applications in the particular context of female behaviour. More, recently Becker (1985) integrates the human capital explanation with the allocation of time in the household context in order to explain the lower hourly earnings experienced by married women. As he points out the increasing returns from specialised human capital provide a powerful force for the division of labour in terms of the allocation of time and the accumulation of work experience within the household. Becker (1985) argues that the allocation of responsibilities (in a household utility maximising framework) to the married female for child care and housework (in which the female is assumed to possess a comparative advantage) has implications for earnings and occupational differences between men and women. A women's assumed comparative advantage in child-rearing and household chores is assumed by Becker to explain the sexual division of labour. This division of labour has clear implications for the accumulation of specific human capital in the labour market and, hence, life-cycle earnings' differentials.

Demand Side Explanations

Theories explaining the existence of a gender wage differential that focus on the demand side also fall into a number of broad categories. Three strands can be identified in the literature,

(i) neo-classical theories of economic discrimination,

(ii) institutional and efficiency wage theories,

(iii) Marxian.

The separate treatment of neo-classical and efficiency wage theories may antagonise some purists. However, since most of the efficiency wage theories to be surveyed here are readily applicable to the concept of the dual labour market it seems more appropriate to group the institutional literature with the efficiency wage literature while bearing in mind the latter's antecedents. There could also be an argument for making no distinction between (ii) and (iii) since the institutional writings have much in common with the Marxian writings. However, though the Marxian writings are only briefly alluded to it was deemed more appropriate to consign them to a distinct category.

2.3 Neo-Classical Theories

It is clear that neo-classical theories of discrimination expressed both in terms of race and gender have commanded significant attention in the literature. In general, the neo-classical theories are set in either a deterministic or a stochastic framework with the former sometimes involving assumptions concerning the structure of the product market *i.e.* whether it is perfectly competitive or monopolistic *etc.* A significant portion of the seminal work in this area again owes its origins to Becker (1971). In Becker's work the abstract concept of prejudice is translated into a taste for discrimination. This taste is rendered economically operational by its introduction as an exogenous argument in a firm's or an individual's utility function. According to Becker's analysis discrimination by race or gender may be employee, employer or consumer motivated. Each of these three channels are examined in turn.

In the context of the consumer based discrimination discriminating consumers are assumed to offer a price $p - d$ for the services of an individual where p reflects the price paid in a discrimination free environment and d measures the intensity of the consumer's taste for discrimination. Consumer motivated discrimination is most likely to impinge upon the earnings of the self-employed and in the service industries where a large proportion of female labour is located. In the long-run and in response to this discrimination one would expect females to either segregate into those occupations characterised by minimum consumer contact or cater exclusively for the female sector.

Discrimination as practiced by employees may also lead to labour force segregation or partial segregation along race or gender lines. The employee discrimination model may again be derived from a utility based framework where employees (or their representative organisations) trade off wages for the ability to work in exclusively male work environments. A logical prediction of the model is the segregation of females into firms' whose employees possess no taste for discrimination. A motivation for the discrimination may come as a consequence of the threat posed to males by the influx of women in terms of job security and earnings. As a consequence trades unions, for example, may in order to protect the interests of their members impose restrictive practices on female entry that are discriminatory. In practice, however, it seems difficult to accept that firms are characterised by the degree of flexibility implied by the above theory[7]. It could, however, be conceded that in the context of race the above employee model of discrimination is more applicable if discrimination is firm or establishment based. There is evidence that this is actually the case for race. If, on the other hand, sex discrimination is occupation based the above analysis is seen to have limitations in its application to gender discrimination.

A more common taste for discrimination model is expressed in terms of an employer's taste for discrimination. The employer's positive taste for discrimination manifests itself in terms of the discriminating firm that sacrifices profit levels in order to indulge in its taste for discrimination. Thus when faced with a wage w for the female it behaves *as if* w(1 + d) is the net wage rate where d again reflects the firms' taste for discrimination.

As Chiplin and Sloane (1976) show this neoclassical model may be expressed more formally. The following assumptions are made:

(i) All firms possess identical utility and production functions.

(ii) Only one commodity is produced and males and females are perfect substitutes.

(iii) The supply of male and female labour are perfectly inelastic.

(iv) The analysis is exclusively short-run with capital given to each firm and output is assumed a variable function of employment levels.

(v) Employers are assumed to maximise a utility function which is described in terms of profit and employment.

[7] Males and females are assumed perfect substitutes in production.

More formally profits for the firm may be expressed as

$$\pi = G(M + F) - w^m M - w^f F \tag{2.1}$$

where π is profits,

G( ) is a strictly concave and increasing function,

M and F are levels of male and female employment respectively,

and w^m and w^f are the equilibrium male and female wages respectively.

The firm's utility function is described in terms of some of the above arguments

$$U = U[\pi, F, M] \tag{2.2}$$

where U is the firm's level of utility.

(2.2) may be re-expressed as :

$$U = U[\, G(M + F) - w^m M - w^f F), M, F] \tag{2.3}$$

Maximising (2.3) with respect to M and F yields

$$\frac{\delta U}{\delta M} = U_\pi(G'_m - w^m) + U_m = 0, \tag{2.4}$$

$$\frac{\delta U}{\delta F} = U_\pi(G'_f - w^f) + U_f = 0, \tag{2.5}$$

and hence

$$U_\pi(G'_m - w^m) + U_m = U_\pi(G'_f - w^f) + U_f \tag{2.6}$$

Under the assumptions that males and females have identical productivity then setting $G' = G'_m = G'_f$ and solving for the marginal productivity condition yields

$$G' = w^m - \frac{U_m}{U_\pi} = w^f - \frac{U_f}{U_\pi} \tag{2.7}$$

Setting $-\frac{U_m}{U_\pi} = d_m$ and $-\frac{U_f}{U_\pi} = d_f$ and substituting into (2.7) yields

$$G' = w^m + d_m = w^f + d_f \tag{2.8}$$

In the absence of discrimination against females (or in favour of males) $U_m = U_f = 0$ and hence $d_m = d_f = 0$. Thus males and females are paid a wage commensurate with their marginal product. Under the assumption of discrimination against females $U_f < 0$ and $d_f > 0$ which implies that the female equilibrium wage is below the marginal product by the amount of the discrimination effect which in each firm is determined by the negative marginal utility to the firm of employing females. Thus in terms of Becker's employer discrimination model with a discrimination coefficient of $d_f > 0$ then $G' = w^m > w^f$. This can be reflected in the market discrimination coefficient which may be expressed as

$$D = \frac{w^m - w^f}{w^f} \tag{2.9}$$

where D is the market discrimination coefficient and w^m and w^f are the equilibrium market wages.

The assumption concerning males and females being perfect substitutes again is deemed relatively unrealistic. The "physical distance" aspect of the taste model could also be deemed more than slightly unrealistic. It could be argued that some firms possess a d of zero for secretaries but a significantly positive d for the professional and skilled occupational categories. Thurow (1975) among others attempts to re-express the problem in terms of "social distance" rather than "physical distance". Nevertheless, as Cain (1986) points out this re-interpretation does not vitiate the main conclusion of the Becker model, that of an equilibrium differential in wages favouring males.

In the perfectly competitive context another issue relates to the variance across industry of the firms' d's. It can be argued that the wider the dispersion in discrimination coefficients the lower is the market discrimination coefficient outlined in (2.9) above. This follows from the fact that most workers who are subjected to discrimination choose those firms who possess low d_i's. Since discriminating firms with high d_i's do not bid for female labour (as a consequence of the negative marginal utility associated with employing females) they end up paying w^m. If most females congregate into firms with low d_i's (or $d_i = 0$) the wage received is closer to their marginal product. Hence the market discrimination coefficient of (2.9) narrows.

A more important aspect of the variance in the discrimination coefficients is in its implications for the long run. The additional profits earned by the non-discriminating firms (assuming males and females are perfect substitutes) in hiring female workers may act as an incentive for these firms to price cut the discriminators in the product market and expand production at the expense of these firms. The demand for labour of the discriminating firms will consequently be reduced. If a constant cost industry structure is assumed to characterise the long-run then the long-run equilibrium outcome will be where firms with low d_i's end up employing all the females and all the males with the latter losing their previously advantaged status. Thus in a perfectly competitive economy characterised by constant costs discrimination cannot persist and in Arrow's (1972) view a logical prediction of the Becker model is the "absence of the phenomenon it was designed to explain".

The "logical" outcome is however contingent on, among other things, a constant cost industry assumption which follows from the fact that factor prices do not alter in response to industry expansion. This assumption may not be altogether tenable. Becker (1971) suggests entrepreneurship as a factor of production that is relatively supply inelastic and with the expansion of industry output (as a consequence of price-undercutting by non-discriminating firms) its price would rise altering the cost structure of the firm. This could well be the same for other factors. If it was the case that the non-discriminating firm's industry was characterised by increasing costs then higher costs would offset profit levels and the opportunities that the non-discriminating firms would have relative to the discriminating firms would be eroded. Thus, in an increasing cost context there exists no guarantees that a discriminating cost differential would persist in order to be exploited to the benefit of the non-discriminating firms.

Other forms of market structure are more amenable to the persistence of discrimination in the long-run. A monopoly in the product market can be cited as an example in this regard. Since there exists only one employer and excess profits in the long-run, no variability in the taste for discrimination and plenty of latitude to sacrifice some profit to indulge the taste for discrimination. However, if a monopolist possesses no power in the input market it cannot influence the wage a male or female labour unit receives. Through d the discriminating behaviour of the monopolist would manifest itself in terms of employment segregation. It is equally possible that the segregated composition of a monopolist's work-force may have little to do with the discriminatory tastes of the monopolist. A monopolist with a $d = 0$ could be consistent with a segregated work-force derived from the fact that monopolists tend to be both capital and skill intensive operations. If through pre-labour market discriminatory processes females fail to gain access to skilled training, then, they will also fail to gain access to jobs in monopolists' firms. Furthermore, unions tend to organise in those industries with higher potential rents[8]. If unions adopt discriminatory procedures against females in their attempts to protect the earnings and job security of their membership, then a non-discriminating monopolist would again be observed with a segregated labour force.

The monopsonist market structure allows for labour market influences to be exerted in the labour market by the monopsonist firm. Monopsonists are the sole buyers of labour in the labour market but may or may not possess any monopolistic influence in the product market. Discrimination in this context may be interpreted in terms of the concept of monopsonistic exploitation. This is measured by $\frac{vmp - w}{w}$ which is interpreted as the mark-up of the value of the marginal product over the equilibrium wage. The discrimination interpretation hinges on a wider monopsonistic exploitation of women relative to men. This is assumed to come about if the female labour supply elasticity is more inelastic than the male's. However, empirical evidence (see Killingsworth (1983)) suggests that female labour supply is more elastic than male labour supply and thus a major tenet of the discrimination thesis is undermined.

Comanor (1973) in an application to racial discrimination views discrimination as a normal consumption good. In the case of an owner-manager higher incomes through profits increase

[8] The Marshallian laws of derived demand may be invoked to explain the existence of greater potential gains in monopolistic rather than in competitive industries.

consumption of the discrimination good. Alchian and Kessel (1962) argue that utility companies in the US operating subject to rate-of-return regulations face an externally imposed profit constraint. This allows managers to indulge in discrimination with no concomitant sacrifice of profits. Ashenfelter and Hannan (1986) highlight a discriminatory explanation in terms of the separation of owners and managers. Under the assumptions of imperfections in the markets for capital and managers and the separation of ownership control, discrimination may be practiced by managers at no pecuniary cost to themselves. The imperfections associated with the capital markets prevents the owners of capital from distinguishing between observed profits and maximum profits. Costly monitoring of the firm in the context of owner-separation provides conditions inside which discrimination as exercised by managers can persist.

The role of trades unions in the labour market may also act as a medium to reduce or re-enforce discriminatory factors. As Sloane (1985) points out a lot depends on the structure of the trades union itself, *i.e.* whether it is organised on an industry-wide basis or on a craft basis. In the former case females may have little problem in joining such unions but may encounter problems in determining its agenda through lack of senior representation. In the latter case females may actually encounter discrimination in terms of their ability to actually join these types of unions. Pre-entry labour market discrimination that exists in terms of female access to skilled training may explain this particular phenomenon.

In the context of the deterministic models outlined above marginal revenue product and marginal factor cost are known with certainty. Phelps (1972) suggests that the existence of imperfect information in the labour market may force employers to maximise expected profits under conditions of uncertainty. As Phelps (1972) points out gaining information about applicants is excessive and costless indices like race or gender may be used by employers as selection criteria.

A central theme of the statistical theory of discrimination is the use of performance indicators by employers to assess the respective productivities of workers. Average wages are paid in terms of group mean performances. If labour force experience is used as a performance indicator, and married women participate less in the labour force than single women, then, through a process of adverse selection the wage paid to women will be based on an average of both's performance. Single women are thus being discriminated against on the basis of the female group. This could be interpreted as a version of group discrimination. Aigner and Cain (1977) rationalise the payment of a lower average wage in terms of employer risk aversion. The lower wage is treated as a compensation to the firm for bearing the risk associated with the unreliability of the performance indicator. This has clear parallels with the insurance market literature where the male/female wage differential could be viewed as a premium paid by the female in order to secure a job[9]. As Cain (1986) argues the statistical theories rest on the differential information conveyed by workers concerning their productivity. Discrimination could, in this form, be viewed as some form of market failure the remedies to which may lie in government policy aimed at the educating, training and licensing of workers.

A variant of the above type of model expressed in terms of signaling theory was developed by Spence (1973). Here workers are assumed to invest in a signal (*e.g.* education) which has no intrinsic value (*i.e.* it is not productivity enhancing) other than to indicate the worker's inate productivity. The signal is costly to obtain and relatively more so for less able workers. The investment in signals by workers is on the basis of their perception of their own ability. Employers are assumed to believe the signals and make wage offers in response to these signals. However, there is no guarantee that the equilibrium allocation of signaling investments is socially efficient. There is a tendency for over investment in signaling and, thus, no guarantee of a unique equilibrium. Spence (1973) shows that given similar amounts of signal investment by two comparable groups the consequent Pareto inefficient allocation leads to a disadvantaged position for the discriminated group.

Stiglitz (1974) has also demonstrated how different equilibria may exist for different groups of workers. Additional assumptions made in the Stiglitz model are that individuals are not perfectly certain as to their characteristics and are risk averse. Firms are assumed not to be able to

---

[9] In fact, Dahlby (1983) employs the concept of adverse selection in an empirical application to the Canadian automobile insurance industry. The author detects evidence of statistical discrimination against young single males relative to young single females and suggests that the premium paid by young females would increase in the absence of discrimination.

observe perfectly the individual's characteristics prior to job placement. Under these and some other assumptions Stiglitz shows that if firms believe that groups delineated by gender or race have different distributions of the imperfectly observable characteristics group membership will be used to determine job placements and hence wages. Finally, Lundberg and Startz (1983) integrate aspects of both the Phelps model and the Spence model and assume that investments in signals are productivity enhancing. Assuming there exists a less reliable signal for one group (females) relative to another then under-investment will occur in it with the obvious consequences for job placement and wages.

It is clear that obvious solutions to discrimination that may occur as a consequence of misperceptions lie in trial work periods for or bond posting by applicants in order to assess more realistically productivity. However, the implementation of this type of procedure may prove prohibitive from the firm's cost point of view.

2.4 Institutional and Efficiency Wage Theories

The deterministic neo-classical models outlined in the last section attempted to explain the phenomenon of discrimination in terms of tastes the effect of which was wage differences and occupational differences for individuals with comparable characteristics. The institutionalist view focuses more on differentials that exist for comparable individuals in terms of their access to jobs. The historical origins of institutionalist writings are found in the demands for social reform that emerged in the early part of the twentieth century. Its re-emergence in the US in the 1960's was also coincidental with a period of dramatic social change and upheaval experienced in that country at that time.

Amsden (1980) suggests institutionalism reaches its fullest expression in addressing questions of occupational segregation by gender and female low pay. The institutionalist theories[10] could be allocated without much difficulty to the broader category of segmented labour market/dual labour market theories (SDLM) and in addressing the issue of <u>job</u> discrimination as

[10] Neo-classical theorists may dispute the use of theory as a term descriptive of institutionalist writings.

opposed to wage discrimination build on the seminal work of Fawcett (1918), Webb (1919) and Edgworth (1922).

Nevertheless, a major limitation of the labour market segmentation literature lies in the fact that it does not provide a single or unique explanation for the existence of wage differentials. This follows from the fact that there exists no consensus as to the appropriate segment upon which to concentrate attention. Segmentation may be on the basis of occupation, industry, firm size, unionisation, the firm's product market or a combination of any or some of the above. The major objective of the segmentation theories are to provide a dichotomy[11] of the labour market on the basis of job characteristics. Some jobs are characterised by low wages, poor working conditions, unstable employment and no on-the-job training. Other jobs are characterised by high wages, good working conditions, stable employment and the provision of various forms of job specific training. As Wachter (1974) points out the important distinction is between "good" and "bad" jobs and not between skilled and unskilled jobs. Skilled workers may be present in the secondary labour market but unable to gain access to the primary sector as a consequence of institutional or discriminatory barriers to entry.

The primary labour market may be viewed as an internal labour market (see Doeringer and Piore (1971)). In this context institutional rules and social custom rather than competitive forces are viewed as the important mechanism for the allocation of wage rates and jobs among primary workers (see Piore (1983))[12]. Furthermore, due to the existence of job rationing in the primary sector the skilled jobs in this sector are unresponsive to the supply of skilled workers. The policy implications that this has for human capital investment is clear. The secondary sector, on the other hand, is assumed described by an alternative wage setting mechanism. The returns to human capital investment are negligible suggesting a relatively flat age-earnings profile. Primary sector rationing is assumed to induce over-crowding in the secondary sector and hence lower productivity. The sector is also assumed characterised by high turnover and an unstable

[11] The dichotomy of jobs is a theoretical and an empirical convenience that may be viewed as an abstract from a more realistic multi-sectoral world.

[12] Williamson, Wachter and Harris (1975) argue that the internal labour market concept need not be viewed in strictly non-economic terms and provide an efficiency rationale for this particular concept.

employer/employee relationship. The poor conditions that characterise the secondary sector are assumed to imbue the workforce with low morale. Absenteeism and poor attendance are assumed rife in this sector.

The theory has much in common with that developed by Thurow and Lucas (1972) (see below) where little emphasis is placed on human capital characteristics and the allocation of jobs is demand determined. The observed gender wage differential is explained by job discrimination and the failure of females through demand side discrimination to access the primary sector jobs. Confinement to the secondary market is assumed to have adverse effects on worker morale and their "taste" for work which is thus rendered endogenous in this framework. Dual labour market theory can be seen as possessing both direct and indirect consequences. The direct effect is in terms of allocating workers to secondary jobs and the indirect effect comes from a workers demoralisation which contributes to the perpetuation of their secondary status.

Albeit from a class perspective Reich, Gordon and Edwards (1973) provide a working definition for a segmented labour market. They see the segmentation of the labour market as a logical response of the capitalist to the homogenisation of the labour force and the threat posed by a unified labour front. It is thus seen as a consequence of the transition from competitive to monopoly capitalism. Reich *et al.* (1973) define labour market segmentation as the " historical process whereby political-economic forces encourage the division of the labour market into separate sub-markets or segments distinguished by different labour market characteristics and behavioural rules". Segmentation can be by race or gender and can cut horizontally or vertically across occupations. The female segments tend to be consistent with the lower wage segments when compared to males[13]. Though this interpretation has a strict Marxian feel about it the definition itself provides a working framework inside which to analyse the SDLM theories.

One SDLM theory, the job competition theory, due to Thurow and Lucas (1972) owes much to the theory of "queues". Job competition is seen to replace wage competition. Technology is assumed to determine the number and type of jobs with social custom and institutional fac-

[13] Segmentation does not imply segregation (see below).

tors determining wages. Screening devices are invoked by employers to allocate workers to jobs and this is a potential channel through which discrimination may occur. Wages are assumed fixed with macroeconomic fluctuations assumed only to alter the length of the "queues". The theory also emphasises the relative insular position of the firm from external factors thus drawing on the much noted internal/external labour market dichotomy.

Zellner (1972) and Bergman (1974) both present models in some sense related to the dual theories. Bergman's model is presented in terms of the "crowding-out" hypothesis. The arguments presented are comparable to those outlined in terms of the dual theories. The immobility that comes as a consequence of the lack of access experienced by females to the higher-paying occupations results in an abundance of females in the lower-paying occupations. This results in a lower marginal productivity in the "crowded-out" sector and hence lower wages.

As Cain (1976) points out the SDLM theories are, in general, "sketchy, vague, and diverse if not internally conflicting". The orthodox neoclassical theory's exclusive emphasis on the budget constraint (income) and opportunity cost (prices) is replaced in SDLM theories by a greater emphasis on institutional and social factors. Cain (1976) concedes that the limitations possessed by the orthodox theory are contrasted with some advantages the SDLM theories possess in the greater emphasis placed on the historical, institutional and qualitative features that characterise the economic system.

The Cain (1976) survey is generally felt to have had an adverse effect on the development of the institutionalist ideas cited above. As pointed out these criticisms were aimed at the lack of theoretical structure attached to the SDLM concepts. However, with the advent of the efficiency wage literature the SDLM school of thought has experienced something of an intellectual renaissance. Katz (1986) provides a comprehensive review of the efficiency wage literature and attention here focuses on those studies that embody explicit application to the dual or segmented labour market concept.

Different variants of the efficiency wage hypothesis can be invoked to rationalise the existence of dual labour markets. The basic tenet of the efficiency wage theory is the positive relationship between an individual's wage and their productivity[14]. Firms are assumed to pay an equilibrium wage above the competitive wage in order to elicit greater workforce effort, reduce shirking, ensure lower turnover costs and higher worker morale. The equilibrium wage can be obtained in a simple profit maximising framework as that wage which minimises wage costs per unit of labour. Thus, the basic efficiency wage theory can be invoked not only to explain real wage rigidity and involuntary unemployment but can also provide explanations as to why workers with identical characteristics are paid differently. High wages and job rationing can arise in that sector where efficiency wage considerations predominate (the primary sector) allowing competitive market forces to govern the rest of the economy (the secondary sector).

Efficiency wage models can be classified into three distinct categories:

(i) sociological norm models (*e.g.* Akerlof (1982)),

(ii) adverse selection models (*e.g.* Malcolmson (1984)),

(iii) shirking models (*e.g.* Bulow and Summers (1986)).

Akerlof (1982) provides an efficiency wage theory expressed in terms of the sociological concept of group norms. In part wages are determined by the norms of workers efforts and may be above the competitive market-clearing wage. Two sectors are delineated on the basis of the existence or non-existence of such wage based norms. However, it's not clear from Akerlof's analysis that there exists any other distinguishing feature of the primary sector relative to the secondary sector above the existence of norms. It could well be the case that loyalty and sentiment to the firm by the worker group is compatible with low-skill jobs and that on the basis of the traditional taxonomy should belong to the secondary sector. The delineation of the labour market into primary and secondary segments requires something more than the existence of wage norms.

Using the principal-agent concept in a two-period model Malcolmson (1984) generates five features of the labour market some of which are consistent with the SDLM concept. Malcolmson (1984) presents a model where contracts with payment are based on the ranking of an employee's

[14] Here the causality is assumed to operate in the opposite direction to that predicted by human capital theory. Higher wages are assumed to induce and not be a consequence of higher productivity.

performance. The employee cannot verify an employer's observation of his performance. Malcolmson (1984) uses these features to generate a model with a hierarchical wage structure, internal promotion, wage rates that rise with seniority and experience more than with productivity, and wage rates attached to jobs rather than individuals. Some of the features outlined above are consistent with the SDLM structures outlined in Doeringer and Piore (1971).

Efficiency wage models based on shirking have also used the principal/agent concept in an application to the segmented labour market literature. Since workers have some discretion over their performance on the job a clear moral hazard problem emerges. In response to this employers may pay an efficiency wage to ensure that the appropriate incentives exist for the worker to avoid shirking. The efficiency wage is most likely to be paid in that sector where the monitoring costs associated with the detection of shirking are high and less likely in the sector where monitoring is relatively costless. Bulow and Summers (1986) argue that such a model has a ready application to the concept of the dual labour market. Primary jobs by their character are difficult to monitor in contrast to secondary jobs which are relatively routine and performance is easily assessed. Since the cost to the firm in the primary sector of workers shirking is likely to be greater than in the secondary sector (since primary jobs are characterised by greater responsibility) there exists a clear incentive for the firm to provide the workforce with the incentives to avoid shirking. From the workers' point of view the existence of an unemployed "queue" of primary workers, while not influencing the wage paid in any sense, acts as a mechanism for discipline (see Shapiro and Stiglitz (1984)). The higher the efficiency wage the greater the cost to the employee of being fired as a consequence of shirking. Higher efficiency wages imply higher productivity and, if the secondary sector only pay the market-clearing wage, a productivity gap exists between the two sectors.

The shirking model can also be used in an attempt to explain the persistence of discrimination. Bulow and Summers (1986) argue that the persistence of discrimination is related to differential turnover propensities. If one group exhibits a higher turnover propensity and hence a shorter horizon on the job it is likely to require greater inducements not to shirk. Thus the higher turnover group is more likely to be confined to the secondary sector. Turnover models have also featured in the efficiency wage literature and place emphasis on the incentives necessary to ensure the long-term attachment of employees to the firm, in particular skilled employees. Since a female's labour force experience is characterised by interruptions profit maximising employers would have no reason to pay them an efficiency wage. Thus the sectoral segregation of females and the wage disadvantaged position of females can be explained by efficiency wage type arguments.

Katz (1986) outlines a basic objection to the dual labour market approach and argues that if secondary workers are as productive as primary workers and are enviable of primary workers then primary wages would be bid down to clear the market. This, however, overlooks a basic feature of the institutionalists' writings that of the "scarring effects" associated with secondary attachment. The low morale experienced by workers attached to this sector, which manifests itself in terms of absenteeism and high turnover, generates characteristics that are unacceptable from the point of view of the primary sector employer.

In a study, separate from the main stream of the efficiency wage arguments outlined above, M[c] Donald and Solow (1985) analyse the effects of business cycle fluctuations on a segmented labour market where the segmentation is in terms of a primary (union) sector and a secondary (competitive) sector. The primary sector wage is not determined as an efficiency wage but a Nash solution in an employer/union bargaining model similar to the model outlined in M[c] Donald and Solow (1981). The model has many dual labour market characteristics including immobility across sectors and a "transitional" pool of unemployed who are assumed to "queue" for primary jobs and accept the first one offered. Employment in the primary sector is shown to be very responsive to exogenous changes in product demand while wages are shown to be more sensitive than employment in the secondary sector. Though this theoretical finding conforms to the stylised facts of the US economy it is slightly at variance with some features of the SDLM model where employment instability best characterises the secondary labour market.

It is clear that recent developments in the efficiency and related wage literature have in some sense diminished the dismissive views expressed by Cain (1976). The development of logically consistent theoretical models that are predicated in terms of equilibrium concepts has allowed a certain rigorous structure to be imposed on a relatively nebulous and atheoretical literature. In so doing some credibility and respectability, certainly in the eyes of neoclassical economists, has been restored.

2.5 Radical and Marxian Theories

Finally and briefly, turning attention to the radical and Marxian theories of discrimination. The general view taken by radical economists is that discrimination is wide-spread and persists. The radical explanation is not derived from a taste based model but rationalises discrimination in terms of the benefits that accrue to the capitalist. The taste parameter does not enter as an argument in the capitalist's utility function. This, of course, is a view at variance with the standard theory outlined by Becker, for instance. The capitalists' strategy is based on a divide and conquer precept. This can be achieved by integrating workforces racially or by gender in order that conflicts are caused within the workforce and their bargaining power is reduced. The segmentation thus implied by the Marxian version of events does not imply segregation. However, the outcome outlined here does not explain the existence of a gender or race wage differential. It could be argued that the payment of a premium to one type of worker over another comparable worker where the distinguishing difference between the two is in race or gender is one method to ensure antagonism within the workforce and thus reduced work-force bargaining strength. Roemer (1979) presents a model based on a profit maximising employer and illustrates the incentives to the employer of employing an integrated workforce (over a segregated one) with the existence of a distinct wage preference for one set of workers over another.

The Gender Wage Gap: An Empirical Survey

As is obvious from the fore-going analysis discrimination theories differ in their focus of attention. Though most theories have implications for the gender wage differential others have implications for the occupational segregation of workers. Since the analysis presented in this study concentrates on the magnitude of the gender wage gap this is reflected in the subsequent empirical survey. Nevertheless, the effects of occupational segregation are not ignored. Indeed, the effects of occupational

In line with the theoretical models outlined in the previous section much effort has been expended on attempting to test both the underlying theories implied by the models and to provide estimates of the gender wage gap holding productivity characteristics constant. The sequential ordering of theories surveyed in the last section is roughly adhered to here. The empirical literature to be surveyed below could be classified into four distinct groups. The first group of studies concentrates on direct and indirect empirical tests of the theoretical models themselves. The second group of studies focuses on methodological and other issues associated with the estimation of the gender wage gap. The third group of studies are presented in order to assess the empirical strength of the SDLM models. The final group provides a miscellany of studies that are not conveniently allocated to the first three groups. For the purposes of this thesis most emphasis is placed on the second and third categories since they more readily capture the context inside which the empirical analysis of the thesis is to be presented.

2.6 Direct and Indirect Tests of Discrimination

As was outlined in the theoretical survey above, consumer discrimination is most likely to impinge upon self-employed earnings. Fuchs (1971) using US data found self-employed earnings for females significantly below comparable male earnings. It was also found that fewer women than men were in this particular category. This latter fact is consistent with Becker's segregation predictions. Nevertheless, it is more likely that certain advantages can accrue to females where there exists consumer contact (see Shephard and Levin (1973)) and that ethnic minorities are more at risk from this form of discrimination. Most recently Kahn and Shaw (1988) provide evidence of black compensation disadvantage in the context of professional basketball in the US which the authors view as being consistent with consumer discrimination.

Employee models of discrimination have been relatively marked by their absence from the literature. The empirical versions of the employee model have concentrated on the relationship between work-force integration and wage compensation. Chiswick (1973) provides a study in terms of race the results of which are consistent with Becker's theory. Chiswick's study was state-based and it would seem more appropriate to express the empirical problem in terms of industry and/or occupation.

A wider empirical literature is available for the employer based discrimination models. The empirical implementation of the employer based model is a test for the existence of an inverse relationship between the female/male net wage ratio and the female/male employment ratio. This may be expressed by the following relative demand curve

$$\frac{w^f}{w^m} = F\left[\frac{E^f}{E^m}\right] \tag{2.10}$$

where w and E refer to net wages and employment levels respectively with the superscripts m and f denoting males and females respectively.

A negative coefficient on the relative employment term is shown to be consistent with the Becker theory. As the wage difference between males and females (who are assumed perfect substitutes) widens the employer's aversion to employing females weakens and more females are recruited. Zabalza and Tzannatos (1985) use such a functional form to assess the effects of anti-discrimination legislation (ADL) in the UK with mixed results. In a number of their specifications positive relative employment coefficients are reported. However, on the introduction of step-wise dummies to account for the cumulative effects of ADL the coefficient becomes statistically significant and negative. The authors provide no interpretation for this result in terms of discrimination theory. The result taken at face value suggests that only after controlling for the cumulative effects of ADL does a discriminatory result emerge which appears counter-intuitive. However, the appropriate level at which to examine the employer model is at the industry level. In general studies of this kind (Ashenfelter (1972)) have provided mixed and unconvincing results.

The indirect tests of the employer based discrimination model usually attempt to posit a significant relationship between employment ratios or wage ratios and proxies for market power such as profits or concentration ratios. The underlying tests are based on the assumption that perfectly competitive firms that discriminate cannot persist in the long-run. Cain (1986) has criticised the implicit assumption in such empirical models that monopoly power in the product market implies monopsony power in the labour market. He also offers criticisms on the use of relative employment variables (as in Oster (1975) or Lukaetich (1979)) as opposed to relative wage variables. The criticism stems from the fact that since the focus of attention is wage discrimination the use of segregation indicators is inappropriate. Though this may be interpreted as a relatively valid criticism the use of such indicators is not outside the purview of Becker's theory. The theory itself has clear implications for employee segregation and such effects can occur in an employer taste for discrimination model. However, the use of relative earnings (as in Chiplin and Sloane (1976)) is clearly inappropriate failing as it does to take into consideration variation in hours *etc*. Of the empirical studies cited in Sloane (1985), (pp.100-101) only Haessel and Palmer (1978) and Luksetich (1979) report a correctly signed significant relationship between discrimination measures and market structure in the context of gender. However, Ashenfelter and Hannan (1986) present results that are consistent with the existence of a negative relationship between market structure defined by market concentration and the firm's employment of women in the US banking system. The use of this firm-specific data overcomes the problem of inter-industry differences that besets some of the other studies cited above (Oster (1975) and Luksetich (1979). Furthermore, Ashenfelter and Hannan (1986) found that an individual bank's share of the market was insignificant in terms of the employment ratio and suggested that the relationship between employment and concentration is due to differences across markets rather than across banks.

In general from the above brief survey of empirical work designed to examine gender discrimination it is clear that no overall picture supportive of the theoretical models emerges.

2.7 Estimating the Gender Wage Gap

Despite the lack of success in defining a robust empirical model with which to test the Becker theories in their various forms no effort has been spared in the empirical attempts aimed at quantifying the magnitude of the gender wage gap[15]. The most enduring methodology that has been used extensively in empirical work on both gender and race differentials is the "index number" approach first suggested by Oaxaca (1973). This approach involves the estimation of

[15] An explanation of the wage gap concept in the union context is provided by Lewis (1986) and in the context employed here is taken to be the excess in the wage for a female if male over the actual female wage at a given specification of characteristics and work conditions *etc*.

separate gender wage equations. Observed female (male) wages and predicted female (male) wages simulated on the assumption of a male (female) wage structure are compared. The difference between the two wages is assumed to provide an estimate of discrimination. Cain (1986) cites twenty US studies over a period of eight years (1973-1981) most of which employed the "index number" approach.

The methodology, however, is neither an ideal nor an accurate measure of the discrimination effect. The discrimination coefficient based on Becker (1971) may be expressed as follows:

$$D = \frac{E(w_s^f) - E(w^f)}{E(w^f)} \tag{2.11}$$

where w_s^f is the simulated female wage on the basis of the male wage structure and w^f is the observed female wage and $E(\ )$ is the expectations operator. The empirical counterpart for the above expression is given by

$$D = \exp(X_i^f \Delta\beta) - 1 \tag{2.12}$$

where a male wage structure is assumed in the absence of discrimination and X_i^f is the set of characteristics for the i^{th} female individual in the sample and $\Delta\beta$ is the difference between the male and female wage coefficients from the OLS estimation of the male and female wage equations. As a summary statistic the discrimination coefficient estimated on the basis of the sample means of the female characteristics is usually reported.

This residual measure may include non-discriminatory elements but may also exclude other discriminatory elements. For example, if female access to certain occupations is restricted as a consequence of discrimination then including occupation controls (through exogenous intercept shifts) may attribute to the explained portion of the observed wage difference a potential discriminatory component. Thus, the inclusion of occupation controls may lower the discrimination estimate. However, their arbitrary exclusion may attribute to the unexplained part of the observed wage difference justified compensating differentials that emerge as a consequence of occupational attachment. The inclusion or exclusion of occupation controls may depend on the objectives of the researcher. Crucial to an accurate measure of the discrimination effect is the correct measurement of all the relevant productivity variables (*e.g* education and labour force experience). The cruder the proxies used for these variables the greater the potential bias in the discrimination coefficient. Furthermore, the omission of key variables that may influence wages (*e.g.* ability or motivation) will also bias the discrimination coefficient estimate.

The "index number" approach takes its name from basic statistical theory and is so called because it involves weighting a given set (or basket) of characteristics by different prices (coefficient estimates). Its construction is analogous to the way the Laspeyres or the Paasche indices are constructed. The calculation of the "index number" is contingent on which coefficients (or weights) are chosen as the base weights. The Laspeyres and the Paasche pose conceptually different questions and this difference follows over into estimating the discrimination effect. The discrimination effect can be calculated by simulating a male wage on the basis of a female wage structure or conversely by simulating a female wage assuming a male structure. Differences between the actual and the simulated wages provide conceptually different estimates for the discrimination effect. Alternatively, a geometric average of both effects can be employed. Sloane (1985) graphically illustrates the "index number" problem in its application to estimating the discrimination effect and shows under certain assumptions how assuming a female wage structure in the absence of discrimination leads to a larger estimate than in assuming a male wage structure. However, *à priori* it is impossible to state where either estimates lies in relation to the other in a multivariate case.

In the empirical literature Oaxaca (1973), Corcoran and Duncan (1979) and Siebert and Sloane (1981) present estimates for both types of calculation. Blinder (1973) and Reilly (1987) assume the male wage structure best characterises the non-discriminating environment and Greenhalgh (1980) in presenting discrimination effects provides a geometric average of both. In the context of the Becker model of employer discrimination outlined above, the assumption of a male wage best describing the discrimination-free environment appears relatively innocuous. All it simply implies in terms of equation (2.8) is $d_m = 0$ *i.e.* no nepotism or favouritism on the part of employers towards males. One short-coming of a large number of the empirical papers that have attempted to calculate the differential is their failure to calculate the associated standard errors without which confidence in the estimates may be diminished. All of the studies cited in the last paragraph with the exception of Reilly (1987) fail to do this. However, the standard errors reported in Reilly (1987) are not based on D but the logarithm of one plus D.

Another important issue to be addressed is the appropriate dependent variable to use in estimating the gender wage difference. Some studies have used annual earnings *e.g.* Greenhalgh (1980), Siebert and Sloane (1981), Lucifora and Reilly (1988) and of Cain's (1986) twenty citations mentioned above over half use some form of income or earnings dependent variable. Though in many instances this is dictated by data availability it seems more appropriate to use net hourly wages since most of the theoretical stories outlined in particular by Becker are described in terms of net wages.

A potentially more important problem is the appropriate explanatory variables and this was alluded to above. Since females, in general, possess a markedly different pattern of labour force behaviour than males an obvious problem emerges in attempting a comparison with males. Their participation rates are lower and their turnover rates are higher. This has clear implications for the calculation of the discrimination coefficient. Polachek (1975) outlines the potential biases that may occur as a consequence of the opposite effects that family characteristics in particular may have on male and female wage structures. In Polachek's view neglecting to take account of these structural differences may lead to an overstatement of the discrimination effect. Greenhalgh (1980) suggests that single male and female workers, who are assumed to possess a similar attitude in the accumulation of human capital investments, provide the appropriate comparator group in quantifying the effects of wage based discrimination.

Nevertheless, this approach may be interpreted as a relatively narrow definition for discrimination. The phenomenon of discrimination can impinge upon married women as much, if not more, than single women. The behaviour of married women must in some way be assessed if an adequate estimate of discrimination is to be provided. The practical problem this creates which is highlighted by inadequate work histories for married women probably explains the empirical focus of attention on single workers. The attendant problem of selectivity bias that follows from the potential existence of some systematic process that governs the observation of women workers in the labour market, *i.e.* women who participate may not be a random sample of women from the population as a whole, also creates econometric problems. Failure to adequately take the selectivity effects into account can lead to biased discrimination coefficients. Few studies have explored how this particular form of bias effects the discrimination estimates. Zabalza and Arrufat (1985) and Dolton and Makepeace (1986 and 1987, a) provide notable exceptions in this regard with the latter studies reporting vast differences between corrected and uncorrected estimates of the discrimination coefficient. It should be said that few studies in general have analysed the effects of any form of selectivity bias on the discrimination coefficient with only Reimers (1983) being another notable exception in the context of race discrimination.

A related question to the above is the manner in which the labour force experience variable is calculated and in particular for the female group. In most studies (Oaxaca (1973), Blinder (1973), Siebert and Sloane (1981) and Greenhalgh (1980)) this is calculated as the difference between an individual's age and years in schooling less six. This is termed a potential experiance variable and in the particular case of married women is likely to gravely over-state labour force experience. Zabalza and Arrufat (1985) and Miller (1987) both show that use of such an incorrectly measured variable attributes more of the observed gender wage gap to the unexplained component leading to an overstatement of the discrimination effect. Zabalza and Arrufat (1985) using Zabalza's (1983) work on labour supply provide estimates of actual female labour force experience and home time. This allows for a more reasonable experience variable and the discrimination estimates are shown to decline when this is used.

Another potentially serious problem alluded to above concerns the appropriate treatment of occupations. Including occupation controls as exogenous intercept shifts in wage equations is designed to control for the existence of compensating differentials across occupations. There are two distinct problems associated with this approach. Firstly, if female access to certain occupations is prohibited by the exercise of an employer's discriminatory power the effects of compensation and discrimination are confused. Secondly, the assumption that occupational level can be treated as an exogenous variable is a relatively strong assumption and should be statistically tested.

In response to the first problem Brown, Moon and Zoloth (1980) provide a modification to the "index number" approach that allows for the wage effects of occupational segregation to be isolated. An occupational attachment model is estimated in conjunction with occupational wage equations. Occupational wage equations are estimated and thus the compensating differential effect is explicitly controlled for allowing a "cleaner" discrimination effect to be isolated. Using US data Brown *et al*. (1980) found that gender wage differences within occupations explained more of the observed gender wage gap than did the allocation of workers across occupations. Miller (1987) using a similar approach produced comparable findings for the UK. The major deficiency in both these studies, however, is their failure to take into consideration the effects of occupational selectivity bias which has clear implications for the discrimination coefficient estimate. This is the second problem outlined above and though a few studies have examined occupational attachment or choice (Schmidt and Strauss (1976) and Boskin (1975)). Less attention has focused on the effects of occupational endogeneity on wages. Hay (1980) and Dolton, Makepeace and Van Der Klaaw (1987)) provide two exceptions in this regard. However, to the author's knowledge no studies to date have examined the issue of occupational endogeneity and its effects on wage discrimination estimates.

2.8 Empirical Applications of the SDLM

The two basic tenets of the SDLM hypotheses are:

(i) the existence of two distinct wage-setting mechanisms in the economy characterised by high and low wages. It's further assumed that the returns to education and labour force experience are greater in the primary sector than in the secondary sector.

(ii) The absence of intersectoral mobility. Secondary workers are assumed unable to cross-over to primary jobs and primary workers are assumed unwilling to take the more abundant secondary jobs even if rendered unemployed. In response to an exogenous change in product demand unemployed primary workers are assumed more likely to join a pool of transitionally unemployed workers and join a "queue" for the rationed primary jobs.

However, the existence of two wage setting mechanisms though consistent with SDLM theories is not inconsistent with the predictions of human capital theory. If an individual can move from one sector to the other then the existence of a sector with no human capital returns is irrelevant. Thus (i) may be regarded as a necessary condition for the existence of dual labour markets with (ii) providing both a necessary and sufficient condition for its existence.

The first hypothesis has been the subject of some empirical investigation. Industries and/or occupations are divided arbitrarily into two sectors on the basis of job characteristics and/or worker characteristics. Statistical tests of the difference between two wage equations are then carried out. Osterman (1975) and Carnoy and Rumberger (1980) find support for the existence of two wage setting mechanisms while Zucker and Rosenstein (1981) find the reverse. The arbitrary classification of industries or occupations into primary and secondary markets on the basis of certain characteristics may partly explain the conflicting set of results obtained. None of the above studies address the econometric issue of sample selectivity bias associated with an individual's sectoral attachment.

The existence of rationed primary jobs has been analysed indirectly through an examination of worker mobility. Leigh (1976) and Schiller (1977) both provide empirical evidence in support of the proposition of mobility thus violating, as they see it, a basic assumption of the SDLM hypothesis. Figart (1987) identifies trades unions as being partially responsible for a gender differential in mobility. The effects of unions are seen as having both a wage effect and an internal labour market effect that operates to the disadvantage of females. Figart also argues that most of the benefits of unionisation accrue to males.

In the particular context of married woman, there is some evidence that occupational behaviour is characterised by a degree of downward mobility. Though Beller's (1985) evidence for the US indicates an increasing female representation in traditional male jobs both Stewart and Greenhalgh (1984) and Dex and Shaw (1986) provide UK evidence of the consignment of women to lower status jobs after labour force interruption. However, the latter authors find little evidence of this for the US. This latter set of results need not be viewed as either refuting or confirming the rationing SDLM hypothesis. Firstly, immobility due to firm specific human capital investments is consistent with human capital theory and mobility within the primary sector does not vitiate the SDLM hypothesis. The downward occupational mobility observed in the UK may be explained as much by human capital depreciation as by the existence of a distinct SDLM with rationing. However, Piore (1980) argues that rationing is most likely to impinge upon women and ethnic minorities during recessions. Employers may impose rationing constraints on the basis of gender and race. Thus in the context of a transitional pool of unemployment employers may use race and gender indicators in the allocation of primary jobs in such a way that women and blacks may lose out especially in a recession.

A clear implication for a gender interpretation of the SDLM hypotheses is the segregation of females into a restricted set of occupations with a consequent depression of the wage. The female intensiveness of a given occupation is likely to impinge upon the wages of males represented in these occupations. If male workers in female intensive occupations experience negative wage effects then this is consistent with what the SDLM advocates predict for the secondary sector. The abundant availability of female labour in these occupations lowers the equilibrium wage. Johnson and Solon (1984) and Lucifora and Reilly (1988) provide empirical evidence for the US and Italy respectively of a well determined negative effect of female occupational intensity on the male wage.

All of the foregoing studies provided indirect tests of the existence of a dual labour market. Dickens and Lang (1985) provide an advance on these indirect tests. Their study also overcomes the classificatory problem by using an econometric model with endogenous switching. This allows the estimation of two sectoral wage equations where the regime governing an individual's wage is unknown *ex ante*. The Dickens and Lang test[16] is based on proportionality between the switching coefficients and the difference between the wage equation coefficients. The rejection of proportionality constitutes a test of rationing in the Dickens and Lang framework.

---

[16] The test assumes income as opposed to utility maximising agents, no costs of mobility across sectors and a known distribution of the wage equation unobservables.

Heckman and Hotz (1986) argue that the rejection of the Dickens and Lang test is uninformative. The rejection may occur for a number of reasons *e.g.* the existence of more than two sectors, the existence of utility maximising rather than income maximising agents *i.e.* certain individuals may prefer the non-pecuniary benefits of secondary sector employment to the income benefits of primary sector employment. The existence of costs of sectoral mobility and false distributional assumptions may also explain the rejection. Furthermore, the likelihood ratio test reported by Dickens and Lang is in error since under the null of no dualism (market clearing) sectoral attachment probabilities are not defined. Thus according to Heckman and Hotz (1986) the dual labour market hypothesis is untestable. However, as Dickens and Lang (1988) argue the Heckman and Hotz criticism is based on an inability to test for market clearing. If this is the case the criticism is equally valid of neo-classical human capital theory.

2.9 Other Studies

Some other studies have attempted to focus attention on the discrimination issue in other ways. Madden (1987) using a survey of displaced workers finds that after displacement (due to an exogenous macroeconomic shock) women experienced a greater wage loss than comparable males. If the human capital explanation is correct women, because they undertake less firm specific investment than men, encounter less of a wage loss after displacement than comparable males. The finding Madden argues is consistent with a gender wage differential that is discrimination related. The smaller set of job opportunities available to women as a consequence of discrimination ensures a smaller likelihood of females securing a job with a wage comparable to their job prior to displacement.

An alternative approach to the estimation of sex discrimination has been proposed by Kamalich and Polachek (1982). In this case the methodology attempts to provide a direct estimate of the effects of sex discrimination while avoiding the biases associated with the "index number" problem. The method used is referred to as "reverse regression" and concentrates on differentials due to job qualifications rather than to earnings or wages. If discrimination exists, then in terms of this analysis, one would expect to find higher mean qualifications among the females for any given wage level. The conclusions of Kamalich and Polachek (1982) was that clear-cut discrimination was not found to exist. However, it should be pointed out that the "reverse regression" approach is not itself free of bias or criticism. Goldberger (1982) has brought the validity of this approach into question and Solon (1983) has demonstrated the inconsistency of the "reverse regression" approach in estimating discrimination effects and argues that it is biased against detecting discrimination.

Chiplin (1981) suggests concentrating less on the estimation of wage or qualifications equations and more on trying to identify a structural demand relationship based on, for instance, firms hiring and promotions decisions. Chiplin adopts an "hedonic offers" framework in an attempt at examining sex discrimination in the context of university selection procedures. No evidence of sex discrimination is found to exist in terms of the data set used by Chiplin (1981) and Dolton (1984) using a similar approach and similar data provides evidence supportive of Chiplin (1981).

2.10 Conclusions

From the above brief theoretical and empirical survey a number of points emerge.

(i) No clear theoretical description of the discrimination phenomenon dominates and this is a view shared by Cain (1986).

(ii) Not surprisingly in the light of (i) none of the empirical tests of discrimination, either direct or indirect, provide convincing evidence one way or the other.

(iii) Some of the econometric issues relevant to the estimation of the gender wage gap have in most cases, but certainly not all, been conveniently ignored.

It is the development of the last point that is most relevance to this thesis. As a means of illustrating (iii) table 2.1 reports discrimination estimates based on six relatively recent UK studies. The estimates reported represent only a subset of some of the cited studies' estimates. The discrimination effects of Greenhalgh (1980), Siebert and Sloane (1981) and Miller (1987) are based on OLS and are unconditional (or reduced form) estimates. Zabalza and Arrufat (1985) and Dolton and Makepeace (1986 and 1987, a) present estimates conditional on the participation decision. The selectivity bias effects originating through participation are corrected for using the Heckman two-step estimator and in the latter authors' case also through a maximum likelihood estimation technique. It is clear from the reported estimates that significant differences attach to the choice of estimator. Miller's (1987) estimates illustrate the problem of using the artificial potential experience variable over the actual experience variable. In using the former variable more of the observed differential in wages is allocated to the residual producing a higher discrimination effect.

In the context of the analysis presented in this thesis neither the issue of potential versus actual experience nor the issue of participation selectivity bias is deemed a major problem. In the former case this follows from the nature of the data set used (to be described in the following chapter) where job information is sufficient to allow the precise calculation of the experience variable. The latter is not deemed problematic since most of the young workers (male and female) are single and the terminal age for inclusion in the sample is twenty-four.

As mentioned in the introductory chapter a major thrust of this thesis is an analysis of an econometric issue that to the author's knowledge has received little attention in the literature. From the econometric point of view the manner in which occupations are treated may be of crucial importance. The standard treatment is to assume exogeneity. However, there is strong evidence to suggest that occupational attachment is endogenous and this has clear implications for the wage equation and discrimination estimates. The correct treatment of occupations is rendered all the more important by evidence provided in Brown *et al.* (1980) and Miller (1987) suggesting a significant portion of the unexplained differential is intra-occupational. A finer occupational discrimination effect can only be obtained by estimating, for example, one digit occupational wage equations and correcting for occupational selectivity bias. Some of the analysis presented below attempts to address this particular issue.

Table 2.1

Discrimination Estimates

Author	Data	Estimate	Method
Greenhalgh (1980)	GHS (1975)	10.0%	OLS
Siebert & Sloane (1981)	Establishment Level Data	0%-16%	OLS
Zabalza & Arrufat (1985)	GHS (1975)	1.3%-12.4%	Heckman
Dolton & Makepeace (1986)	Graduate Survey (1977)	19.8%	OLS
		14.4%	Heckman
		12.4%	MLE
Dolton & Makepeace (1987, a)	Graduate Survey (1977)	7.6%	OLS
		4.2%	Heckman
		5.9%	MLE
Miller (1987)	GHS (1980)	33.6%	OLS
		16.4%	OLS

GHS is the General Household Survey. All the above estimates are based on assuming a male wage structure in the absence of discrimination with the exception of Greenhalgh (1980) who calculates a geometric mean estimate based on male and female wage structures. Heckman is the Heckman two-step estimator described in Heckman (1979) and MLE is maximum likelihood estimation. The selectivity bias corrected for in Dolton & Makepeace (1986 and 1987, a) is for participation. The first estimate reported for Miller (1987) is based on using potential experience for females the latter on actual experience calculated as in Zabalza (1983).

Chapter Three

Description of the Data Set and Background

3.1 Background

The purpose of this chapter is to provide a description of the data set proposed for use in the empirical analysis. Prior to this, however, it may be of some interest to set the overall labour market context inside which this analysis should be placed. Throughout its recent history upto the early 1960's the Irish demographic experience has been one of emigration and declining population. The lowest ever population in the history of the state was recorded in the 1961 census with a population of 2,818,000. Though a steady increase was achieved upto the 1971 census (2,978,000) the most notable increase occurred upto 1981 when the population reached 3,443,000. By any standards this rate of increase was exceptionally high with a proportionate population increase of 15.6% achieved between these two census years. This compares to the European group of OECD countries' proportionate growth rate of 6% over the same period.

Table 3.1 illustrates the major determinant of this particular phenomenon. Between 1951 and 1971 high rates of natural increase were off-set by even higher rates of net migration outflows. Only in the 1971-1981 period was this trend reversed largely due to the net external inflow of population in those inter-censal years. Table 3.2 shows the percentage change in population by age group with the 15-24 age group characterised by the second largest increase over the 1971/81 period. It is clear from both these tables that between the years 1971 and 1981 a dramatic change in the pattern of net migration occurred. A major part of this change is attributable to the substantial reduction in the migration of young workers in the 15-29 age group allied to a net inflow of migrants in the 30-44 age group (mostly coming from the UK with young families). Together both these factors contribute to a population structure in 1981 where nearly 50% are under the age of 25 years. The corresponding OECD Europe average is 38%.

Using the 1981 Census of Population Table 3.3 outlines the main features of the Irish labour market contrasting the youth and adult labour markets. The dramatic feature of table 3 is the sizable proportion of the workforce under the age of twenty-five. Nearly 40% of those at work are in this less than twenty-five category, a fact that provides adequate motivation for concentrating attention on this particular segment of the labour market.

Table 3.1.

Population Change in Inter-Censal Periods.

Period	Change in Pop.	Natural Increase	Net Migration
1951-1961	-142,300	+266,500	-408,800
1961-1971	+159,900	+294,500	-134,500
1971-1981	+465,000	+361,000	+104,000

Source: Censuses of Population of Ireland (1951,1961,1971,1981).

Table 3.2.

Population Change in Inter-Censal Years by Age Groups.

Age-Group	% Change(1961/71)	% Change(1971/81)
0-14	+6.1	+11.5
15-24	+23.2	+26.0
25-44	-1.4	+34.0
45-64	+1.5	-3.0
≥65	+4.7	+11.6
Total	+5.7	+15.6

Source: Censuses of Population of Ireland (1951, 1961, 1971 and 1981).

Table 3.3.

Main Features of the Irish Labour Market in 1981.

Category	15-24	≥ 25	Total
(a) At Work	316.1	834.6	1150.7
(b) Unemployed	33.9	79.0	112.9
(c) Seeking First Job	18.1	1.6	19.7
(d) Education	199.4	3.7	203.1
(e) Other Activities	40.9	878.3	919.2
(f)=(a)+(b)+(c) Labour Force	368.1	915.2	1283.3
(g)=(f)+(d)+(e) Total	608.4	1797.2	2405.6

Source: Census of Population of Ireland (1981).

3.2 Data Set Description

Attention now turns to a description of the sample survey itself. The full title of the survey study commissioned by the EEC and carried out by the Economic and Social Research Institute (ESRI), Dublin, is

" A Survey of Youth Employment and the Transition from Education to Working Life ".

A more comprehensive description of which appears in Sexton *et al.* (1982).

The target group covered in the survey was persons aged between 15 and 24 years of age. All those surveyed had left full-time education and were either gainfully employed or seeking work. No readily available sampling frame exists for this target group. However, the EEC Consumer Attitudes Survey (EECCAS)[17] was used to obtain 22,648 households. Each of these household were visited in three rounds between May 1981 and January 1982. This overall sample of households were surveyed yielding the 5930 individuals in the target age group of 15 - 24

[17] This was a three yearly study conducted in Ireland by the Economic and Social Research Institute and An Foras Taluntais (the Irish Agricultural Institute).

years. The EECCAS, which uses the electoral register as the sampling frame, has a tendency to under-represent the smaller households (*i.e.* households with a small number of electors). The survey of the target group itself was carried out between the months of March and June of 1982. A similar study, also under the auspices of the EEC, was carried out for the Netherlands at about the same time.

Since the interviews were carried out in terms of households, the questionnaires consisted of three main parts:

(I) A Household Sheet,

(II) A Person Questionnaire,

(III) A Job Sheet.

The first part, (I), was concerned with eliciting information about household composition in terms of sex, age, relationship to head of household *etc.* for all the members of the household.

The second part of the questionnaire, (II), deals with questions relevant to the targeted individuals in the 15 - 24 age-group. The data here provide information on

(i) demographic characteristics (*i.e.* sex, age, marital status and nationality),

(ii) details on educational attainment (*i.e.* nature and extent of formal full-time education *etc.*),

(iii) present training (nature and duration),

(iv) basic training (nature and duration),

(v) further training (nature and duration),

(vi) details on the respondent's parents (*i.e.* their economic status *etc.*),

(vii) details on the respondent's own economic status *i.e.* at work or unemployed.

The final part of the questionnaire (III) concerns itself with information relevant to the individual's job and the characteristics of the job. The information contained here details

(i) Date Commenced,

(i) Occupation,

(ii) Industry,

(iii) Location,

(iv) Hours worked,

(v) Gross and Net Earnings,

(vii) If Promoted,

(viii) Date Left.

The individual in the target group was responsible for filling in the personal questionnaire. Since some of the questions asked could be interpreted as imposing a strain on the interviewee's memory, the accuracy of some of the responses may be brought into question. The problem that poor recall poses in eliciting an accurate response to a variable like pay, for example, cannot be understated. The problem of poor recall was to some extent mitigated by the fact that interviewers asked to see documentary evidence of pay in the form of recent pay slips, for example. Though this particular problem can be minimised it can't be eliminated completely and this should be borne in mind in interpreting the subsequent empirical analysis.

Though the information contained in the sample survey is comprehensive it should be pointed out that in terms of part (II) of this questionnaire no information is available on household income or on unemployment benefits received by those in the target group who are unemployed. This clearly imposes some limitations on the scope of the empirical analysis. Though most of the target group sampled live at home and thus would be ineligible for unemployment benefit or assistance the lack of information concerning household income is a limitation especially in the context of examining occupational attachment. Efforts to obtain tapes for the Irish Household Budget Survey of 1980 in order to estimate household income on the basis of the household characteristics of (I) proved unsuccessful as permission to access was only granted under restrictive conditions by the Central Statistics Office in Dublin.

Since the empirical chapters four to seven do not use the identical set of variables the number of observations for each study differs.[18] Thus, each empirical chapter contains a descrip-

[18] This is due to to missing values.

tion of the variables used in estimation in these chapters and appendix A1 contains summary statistics for the relevant variables used in these chapters.

Chapter Four

Gender Wage Discrimination with Exogenous Occupations

4.1 Introduction

The purpose of this chapter is to analyse some of the determinants of wages at the individual level in the Irish labour market for young workers. As pointed out in chapter one this has been the focus of no attention to date in the Irish context. Even for the adult labour market in Ireland little research effort has been expended on analysing the determinants of adult earnings which in large part is explained by either the unavailability or lack of access to suitable data sets. The only exception in this regard is Walsh and Whelan (1976) who used a sample of redundant adult workers to estimate human capital earnings functions in an attempt to analyse, among other things, gender differentials in earnings.

The estimation of such equations provides an important insight into some relevant policy issues. For example, the returns to work experience and educational qualifications can be assessed and the extent to which they differ across gender groups can also be evaluated. A major focus of attention in terms of this chapter is the magnitude of the gender wage differential and to what extent it is explained by differing characteristics and/or by differing returns given the same characteristics. Under certain assumptions the latter case has been interpreted in the literature as providing an estimate of gender wage discrimination.

The major objectives of this chapter are, firstly, the estimation of individual level reduced form wage equations by gender which allows inferences to be made regarding returns to labour force experience, educational/vocational qualifications. Within this framework the effects of occupation, industry and region can also be assessed. The second and more important objective is the estimation of a mean discrimination effect and an examination of how the discrimination effect varies across individual characteristics.

The following section outlines the methodology employed. Section 4.3 describes the data set used and section 4.4, the estimation. Results based on estimating the reduced form wage equations appear in section 4.5 and section 4.6 provides some statistical tests of the underlying specifications. The gender discrimination estimates are obtained in section 4.7 with conclusions presented in 4.8.

4.2 Methodology

The reduced form wage equations to be estimated are assumed to be derived from some underlying structural model of labour supply and demand. The widely applied reduced form Mincerian earnings function (see Mincer (1974)) may be expressed as a function of schooling, post-schooling (or on-the-job) investment, and a vector of personal characteristics. Algebraically, it may be given by (4.1) as:

$$\ln(W_i) = \alpha_i + \beta_1 S_i + \beta_2 X_i + \sum_{j=3} \beta_j Z_{ij} \qquad (4.1)$$

where

W_i is the i^{th} individual's wage,

S_i is years of schooling,

X_i is labour force experience and

Z_{ij} is a vector of personal characteristics (*e.g.* region of residence, occupation and industry *etc.)*.

The theory predicts that in a long-run equilibrium with perfect labour mobility and perfect competition in the labour market, wages reflect an individual's characteristics in terms of schooling, post-schooling investments and compensating differentials commensurate with differing job characteristics. In the context of firm specific training, human capital theory predicts short-run departures from the above equilibrium condition. Individuals who obtain such training are awarded a remuneration initially above their marginal product during training and below it after its completion in order that firms may recoup their outlayed training costs. Because of this one might not expect young workers' wages to be exclusively based on productivity variables particularly where firm specific training is involved.

Discriminatory practices can also be interpreted as inducing departures from a competitive equilibrium. As long as rigidities exist in such a way that the different sexes are paid different market prices for comparable productivity characteristics, then, in the long-run, females working in discriminating firms will be paid less than their marginal product. These discriminatory practices may take the form of employer discrimination, employee discrimination (through monopolistic organisations *e.g.* trades unions) or consumer discrimination as explained in chapter two.

There exists an extensive literature on the calculation of the gender wage discrimination effect. The standard methodology employed was first suggested by Oaxaca (1973) and exploits the properties of index numbers[19]. The "index number" approach involves the estimation of separate gender wage or earnings equations. The calculation of the residual or unexplained difference between the two equations is assumed to provide an estimate of discrimination. However, since the effects of factors omitted from the equation are assigned to the residual its interpretation as an accurate measure of wage discrimination may be inappropriate. Furthermore, since the estimated wage equations are, in general, reduced form it is impossible to disentangle demand and supply side effects. Since discrimination is ultimately a demand side phenomenon the interpretation of the residual as a discrimination effect relies on the strong assumption of supply side neutrality. The assignment of omitted factors and measurement error to the residual renders it impossible *à priori* to establish whether the residual estimate reflects an upper bound or a lower bound discrimination estimate. Despite these caveats the "index number" methodology has had an extensive application in the gender wage discrimination literature. The caveats outlined above should be borne in mind, however, when interpreting the results presented below.

Slightly modifying Becker (1971) the discrimination coefficient for each female[20] individual in the sample or population may be expressed as:

---

[19] Cain (1986) in his discrimination survey cites twenty US studies of race and gender discrimination that use this particular methodology.

[20] The D_i's could also be validly interpreted for the males in the sample or population. However, under the assumptions to be subsequently laid out in this chapter this would by definition be zero for each male in the sample or population.

$$D_i = \frac{W_i^n - W_i^f}{W_i^f} \tag{4.2}$$

where

D_i is the i^{th} female's discrimination coefficient,

W_i^n is the i^{th} female's non-discriminatory wage and

W_i^f is the i^{th} female's actual wage.

The female wage/ earnings equation may be expressed as:

$$W_i^f = e^{X_i^f \beta^f} \tag{4.3}$$

The corresponding male wage/earnings equation is given by:

$$W_i^m = e^{X_i^m \beta^m} \tag{4.4}$$

where the X_i's are vectors of conditioning variables for each i individual and β is a vector of unknown parameters with superscripts f and m denoting female and male respectively.

Re-writing (4.2) gives:

$$1 + D_i = \frac{W_i^n}{W_i^f} \tag{4.5}$$

If in the absence of discrimination a male wage or earnings structure is assumed,[21] W_i^n may be simulated for each female in the population or sample on the basis of the β^m coefficients and the female realisations of the X vector of characteristics. This simulated wage can under the assumption set out above be interpreted as a non-discriminatory female wage. Taking logs of (4.5) and

[21] The "index number" problem emerges here. As is well known the discrimination estimate is contingent on what structure is assumed to exist in the absence of discrimination. If one assumed a female structure, for example, a different estimate would emerge. This is further explored below.

using (4.3) and (4.4) :

$$\ln(1 + D_i) = X_i'\beta^m - X_i'\beta^f \tag{4.6}$$

or more compactly,

$$\ln(1 + D_i) = X_i'\Delta\beta \tag{4.7}$$

where $\Delta\beta = \beta^m - \beta^f$ Thus for each individual in the population or sample a discrimination coefficient can be estimated. For the males this, will of course, be zero but for the females the magnitude of the discrimination effect is determined by the differential in payment for a given set of characteristics.

The vector of β coefficients for both males and females can be estimated by OLS and as a summary statistic the discrimination coefficient can be evaluated at the means of the data. Thus (4.7) may be re-expressed as

$$\frac{1}{N}\sum\ln(1 + D_i) = \frac{1}{N}\sum X_i'\Delta\beta \tag{4.8}$$

where N is the number of individuals in the sample. The empirical counterpart for the mean expression can now be expressed as:

$$\hat{\bar{\psi}} = \bar{X}'\Delta\hat{\beta} \tag{4.9}$$

where $\Delta\hat{\beta}$ is the difference between the male and female OLS coefficient estimates[22]. The calculation of the asymptotic standard error follows Stewart (1987) where the variance may be expressed as:

$$\mathrm{var}(\hat{\bar{\psi}}) = \bar{X}^{f}V\bar{X}^{f\prime}$$

where

[22] $\bar{D}$, the mean discrimination coefficient, may be interpreted as $e^{\bar{X}'\Delta\hat{\beta}} - 1$.

$$V = \mathrm{var}(\hat{\beta}^m) + \mathrm{var}(\hat{\beta}^f).$$

Since $\hat{\beta}^m$ and $\hat{\beta}^f$ are estimated from separate non-overlapping samples no covariance exists between. The asymptotic standard error is thus given by $\sqrt{\mathrm{var}(\bar{\psi})}$. The resultant test statistic is asymptotically normally distributed[23].

The remaining component of the gender wage gap could be interpreted as that part explained by differing endowments. The two components are tied together in (4.10).

$$\Delta\ln(\bar{W}) = \Delta\bar{X}\hat{\beta}^m + \bar{X}^f\Delta\hat{\beta} \qquad (4.10)$$

where $\Delta\bar{X} = \bar{X}^m - \bar{X}^f$ Thus the gender differential in the mean logarithm of wages can be decomposed into that part explained by the possession of differing endowments[24] and that part explained by the payment of different prices for the same endownments. This latter part is, of course, the $\bar{\psi}$ of (4.9).

Any genuine attempt at calculating the effects of sex discrimination should focus on those males and females who could best be described as possessing a similarity in labour market attachment. Greenhalgh (1980) attempts to estimate the magnitude of gender discrimination by reference to the unexplained differential between single men and single women. The underlying assumption being made here is that both these groups are similar in terms of their labour market attachment and hence in terms of their human capital investment. Though this may be interpreted as a relatively narrow wage discrimination definition a similar assumption to that of Greenhalgh (1980) is followed here and in the subsequent empirical chapters.

A number of problems are associated with the use of the "index number" approach in calculating the effects of gender discrimination. Notable among these is the "index number" problem. This stems from the fact that the results based on the above approach are contingent on which

[23] Thus the variances are calculated for the logarithm of one plus the discrimination coefficient and not for the discrimination coefficient itself. Furthermore, standard errors can be calculated for each individual in the sample.

[24] The variance for this expression can be easily calculated as $\Delta\bar{X}V\Delta\bar{X}'$ where V is given by $\mathrm{var}(\hat{\beta}^m)$.

wage structure is initially assumed in the absence of discrimination. The discrimination estimates reported here are based on an assumed male wage structure. This may be regarded as a relatively innocuous assumption since it could plausibly be argued that males are paid a wage that is non-discriminatory. The deterministic model outlined in chapter two could be invoked to support the validity of this particular assumption.

As briefly alluded to above another problem lies in the fact that the estimated equations are reduced form. The discrimination estimate is taken as the unexplained differential that emerges after controlling for the standard set of productivity and other relevant variables. The interpretation of the unexplained residual differential as the pure discrimination effect is only valid if supply side factors *e.g.* female preferences for low paying occupations *etc.* do not impinge upon the unexplained differential. In the context of a reduced form equation it proves difficult to disentangle the demand side discrimination from supply side preferences. The interpretation is also contingent on a correctly specified equation and no measurement error in the explanatory variables that enter that equation, since, as surveyed in chapter two the effects of omitted variables and measurement error are assigned to the residual with a consequent effect on the discrimination coefficient.

Since the focus of this paper is wage discrimination, defined as differences in pay given occupations, the potential discriminatory effect that arises through occupational segregation is ignored. This may be viewed as an unsatisfactory approach assuming, as it does, that there exists no discriminatory restrictions on access to certain occupations. It might be argued that if most gender discrimination is occupation based, then, the above framework is clearly lacking. Brown *et al.* (1980) argue that if the same characteristics that determine wages also determine occupation, then, the "index number" approach may be viewed as an appropriate methodology. In general, they argue that this is not likely to be the case and propose controlling for occupations by incorporating an occupational attainment model into the analysis of wage differentials as a method of overcoming this problem. This particular issue is more extensively explored in chapter six.

For the purposes of this study two "second best" alternatives are presented. Firstly, occupations and industries are controlled for in the standard manner through the use of dummy variables. This approach implicitly treats the occupational distribution of males and females as exogenous and ignores the potential endogeneity of occupational attainment. Regardless of whether the potential endogeneity is a consequence of sample selection on the part of employers or self-selection on the part of workers a potential bias is likely in the estimates. Furthermore, the inclusion of occupation dummies will also lead to an understatement of the returns to formal educational qualifications since access to certain occupations is contingent on possessing certain formal qualifications. Though this simplistic treatment of occupations and industries is accepted as unsatisfactory a failure to control for occupations or industries leads to a confusion between the effects of discrimination and those of a compensating differential.

The second alternative is to estimate the wage equations excluding the occupation and industry dummies. The potential endogeneity problem associated with occupations and industries is thus by-passed in a rather *ad hoc* manner. The consequent estimates of sex discrimination are overstated but the returns to the formal educational qualifications are not under-stated as is the case when occupations are included[25]. Therefore, the results of section 4.5 contain estimates obtained from models that both include and exclude occupation and industry controls.

4.3 Data

The data used in this study are obtained from an EEC commissioned survey carried out by the Economic and Social Research Institute in 1982 and titled " Youth Employment and the Transition from Education to Working Life". The target group in the survey were males and females between the ages of 15 and 24 who had left full-time education and were either actively engaged in employment or actively searching for work. The sub-sample employed in this analysis is composed of those single individuals in the survey who defined their main economic activity as either working for payment or profit in non-agricultural activities. The total number of such cases for which no missing values are present is 1022 (449 males and 573 females).

[25] However, if educational qualifications are chosen endogenously there is latitude for selectivity bias in this regard and the returns to educational qualifications (with or without occupation and industry dummies) are potentially biased.

The variables used are:

Wage: The natural logarithm of the average net hourly wage ($\ln(W_i)$). This variable excludes abnormal hours worked and over-time payments.

Labour force experience: This was calculated as the aggregate duration of all jobs (including the current job) held by the i^{th} individual. The unit of measurement used for the experience variable is, in this case, years. Since this is a precisely calculated variable the usual errors-in-variables problem that characterises similar studies does not exist in regard to this particular variable.

Educational Qualifications:Three (0,1) dummies for the highest educational qualifications obtained by the individual. These qualifications are the

(a) Intermediate Certificate,

(b) Leaving Certificate,

(c) University Degree or its equivalent.

Individuals who have commercial course diplomas are allocated to either the intermediate or the leaving certificate category depending on whether they possess one or other of these public examinations. In terms of estimation the obvious reference group in this case are those individuals in the sample who have no such qualifications.

Vocational Qualifications:

(a) A (0,1) group certificate dummy variable adopting a value of 1 if the individual possesses such a certificate, 0 otherwise.

(b) A (0,1) apprenticeship dummy variable adopting a value of 1 if the individual successfully completed an apprenticeship scheme, 0 otherwise.

(c) A (0,1) basic training qualification dummy variable adopting a value of 1 if the individual successfully completed such a qualification, 0 otherwise. These training qualifications refer to formal instruction outside the context of the educational system leading to formal qualifications.

Promotion: A (0,1) promotion dummy adopting a value of 1 if the individual received promotion on the current job, 0 otherwise.

Firm Size: Four (0,1) dummy variables designed to capture the effects of variation in firm or organisation size. The four categories are firms with less than fifteen individuals (including the proprietor), firms with fifteen and over but less than fifty individuals (including the proprietor), firms with fifty and over but less than a hundred individuals (including the proprietor), and firms with a hundred or over individuals (including the proprietor). The dummy variable adopts a value of 1 if the individual falls into any of these mutually exclusive categories, 0 otherwise. The reference group in terms of this set of dummy variables is the first category of less than fifteen individuals. Obviously, the coding of this variable relates only to those individuals at work.

Urban:A (0,1) urban dummy variable adopting a value of 1 if the individual resides in a town of 1000 or more, 0 otherwise.

Industry: Eight (0,1) industry dummies which are:

(a) Energy, water, extraction and processing of non-energy-producing minerals and derived products and chemicals.

(b) Metal manufacturing (mechanical, electrical and instrument engineering).

(c) Other manufacturing industries (food, drink, tobacco, leather, footwear and clothing, textiles, timber and wooden furniture, paper and paper products, printing and publishing, rubber and plastics).

(d) Building and civil engineering.

(e) Distributive trades, hotels, catering and repairs.

(f) Transport and communication.

(g) Banking and finance, insurance, business services and renting.

(h) Other services (public administration, education, research and development, medical and other health services, government recreational services and personal services).

In terms of estimation the reference group for this set of dummies is the distributive trades industry group.

Occupation: Eight (0,1) occupational dummies which are:

(a) Professional,

(b) Self-Employed (employs others) and Managers,

(c) Salaried employees, *i.e.* insurance and financial agents, auctioneers and valuers, ships' officers and pilots *etc*.

(d) Intermediate non-manual,

(e) Other non-manual,

(f) Skilled manual workers,

(g) Semi-skilled manual workers,

(h) Unskilled.

In terms of estimation the reference group for this set of dummies is the unskilled category.

Region: Five regional dummies which are:

(a) Dublin County.

(b) North-West (Counties Sligo, Mayo, Roscommon, Donegal, Leitrim, Monaghan, Cavan and Galway).

(c) Southern (Counties Cork, Waterford, Limerick, Kerry, Tipperary and Clare).

(d) Midlands (Counties Carlow, Laois, Kilkenny, Kildare, Westmeath and Offaly).

(e) Rest of Leinster (Counties Wicklow, Wexford and Louth).

In terms of estimation the reference group for this set of dummies is the North-West region.

Appendix A1 provides information on the means of the continuous variables and the proportions of individuals in the relevant binary variable categories *etc*. for males and females.

4.4 Estimation

Linear equations are estimated separately for both single males and females in the sample. In the estimation of standard wage functions quadratic terms in labour force experience are introduced allowing for the effects of such experience to vary over its own range. In the context of the labour market for young workers it may be unreasonable to impose *à priori* such a restrictive functional form on the data. In view of this, linear splines (following Stewart and Wallis (1981), pp.201-204) are constructed from the labour force experience variable. This allows for returns to experience to vary over different ranges of an individual's labour force experience. The choice of the appropriate nodes for the linear splines was determined by using standard F-tests. These tests suggest for the male equations the inclusion of two linear splines, one for less than or equal to five years of experience, the other for more than five years of experience[26].

In general, the estimating equation may be expressed as[27]:

$$\ln(W_i) = \alpha_i + \beta_1 \mathrm{EXP1} + \beta_2 \mathrm{EXP2} + \sum_{j=3}^{n} \beta_j \mathrm{DUMMIES} + u_i \tag{4.11}$$

where

$\ln(W_i)$ is the natural logarithm of the net hourly wage defined above. EXP1 and EXP2 are the linear splines in experience. The labour force experience variable is defined as above and is precisely calculated in contrast to the proxy constructs used in Oaxaca (1973) and Greenhalgh (1980) which could be interpreted as possessing a greater potential for measurement error. DUMMIES represents the controlling educational qualifications, firm size, promotion, occupational, industrial and region of residence dummies and u_i is an error term for which assumptions are made below.

Alternatively, experience may be expressed in terms of (0,1) dummies. Eight such dummies are also constructed from the experience variable for each of the eight potential years of labour force experience. Though the results of this analysis are not recorded extensively here, this alternative specification is used for purposes of comparison in terms of the diagnostics of section 4.6.

---

[26] The null hypothesis in this case is the male equation strictly linear in experience. This null is rejected in favour of a piece-wise linear alternative having a five year split by an F statistic of 5.256 distributed as F(1,417). The associated critical value is 3.840 at the 5% level of significance. For the female equation the null hypothesis of linearity in experience cannot be rejected in favour of piece-wise linearity with a five year split. The computed F(1,541) is 3.170. The validity of choosing a similar cut-off point for both males and females may, therefore, be justifiably questioned. However, in terms of this analysis, it is necessary for obvious reasons to estimate comparable male and female equations.

[27] Since estimation is based on those individuals who are currently working another potential selectivity bias problem should be noted. However, this was not found to be a problem as one would expect in a sample of young single workers.

Returning to equation (4.11) it should be noted that the standard assumptions are made concerning the error terms:

(i) $E(u_i) = 0$ (an assumed mean of zero).

(ii) $E(u_i^2) = \sigma^2$ (constant variance).

(iii) $E(u_i X_i) = 0$ (the explanatory variables are assumed to be exogenous and omitted variables are assumed orthogonal to the included ones).

(iv) The parameters are assumed constant.

(v) u_i is assumed normally distributed.

The violation of the first assumption may occur through equation mis-specification thus inducing bias in the coefficient estimates. Since these coefficient estimates are used in quantifying the magnitude of discrimination the correct specification of the equation is of crucial importance. Biases may lead to inappropriate conclusions in regard to the nature and magnitude of any discrimination element.

A priori one would expect violation of the second assumption. In the presence of heteroscedasticity OLS provides unbiased coefficient estimates but a biased and inconsistent covariance matrix estimator. Since consistent standard errors are required for hypothesis testing White's (1980) heteroscedasticity-consistent covariance matrix estimator is employed. The White estimator possesses the appropriate asymptotic properties for hypothesis testing.

Departure from the assumption of normality must also be viewed with some concern. Such a departure may be indicative of some form of mis-specification and may also vitiate the use of standard statistical tests (*e.g.* the t-statistics and the F-statistics *etc.*) which are based on the normal distribution. Nevertheless, the OLS estimator is relatively robust to departures from normality.

In view of the foregoing and the need for some confidence in the results, section 4.6 contains a number of diagnostic tests that statistically test for model mis-specification, heteroscedasticity and the underlying assumption of normality.

4.5 Results

Tables 4.1 to 4.4 contain the male and female regression estimates for the single individuals based on equation (4.11). The first two tables contain estimates based on the inclusion of occupation and industry dummies. The latter two tables contain estimates based on the exclusion of these controls. For purposes of comparison OLS and White-adjusted standard errors are reported in these tables with inference based only on the White standard errors.

Since the dependent variable is in logarithmic form and the explanatory variables are either in levels (the splines) or expressed as (0,1) dummies, some caution must be exercised in their respective interpretations. The coefficients of the experience variables (the splines) are defined as the proportional returns to having an additional year of labour force experience. The dummy coefficients possess a slightly different interpretation. The estimates, themselves, give the differential effect of being in the included category relative to the excluded category (*i.e.* the reference category). Since the dependent variable is expressed in logarithmic terms, the β_i dummy coefficient is interpreted as e^{β_i}.

As tables 4.1 and 4.2 indicate, the returns to labour force experience do indeed vary over an individual's work experience. This is confirmed by examining the coefficients on the splines in both equations and testing to ascertain whether they are statistically different from each other. For both equations this is found to be the case[28]. The actual annual returns to labour force experience are found to be higher for the males in the first five years, 7.5% as compared to 6.3% for the females. For the subsequent years the fall is more dramatic for the males declining to around 1% on average for the years after the fifth year of experience. The decline for the females is less dramatic. They record returns of 2.5% on average for each year subsequent to the fifth. In both cases, however, the higher returns to on the job training in the early years (leading to a steeper wage profile in the initial years of experience) is a finding compatible with the predictions of human capital theory.

[28] The t-statistic associated with the male equation is 3.97 *** and that for the female equation 2.35 ** (where *** denotes 1% and ** denotes 5% significance).

The educational dummy coefficients allow for some conclusions to be drawn concerning returns to educational qualifications. It should be borne in mind that the inclusion of the occupational dummies leads to an understatement of the returns to the formal educational qualifications. It is clear from the results of tables 4.1 and 4.2 that possessing an Intermediate Certificate as one's highest qualification has an effect that is not significantly different from zero for either males or females. The returns to a Leaving Certificate are on average higher for males than for females, however, in both equations the coefficients are neither well determined nor statistically significant from zero. On the other hand, returns to a university degree or its equivalent are quite large and in the case of the males statistically significant. In this case the returns are of the order of 33.6% as compared to the females' 24.8%. These relatively large effects should be treated with some caution given the relatively small number of individuals in each category. In general, therefore, males appear to gain more from educational qualifications in relation to unqualified males than do females in relation to their unqualified counterparts. This is in direct contrast to the Greenhalgh (1980) findings in regard to returns to educational qualifications. In a sub sample of single males and single females under the age of thirty derived from the 1975 UK General Household Survey (GHS) Greenhalgh found that the returns to educational qualifications were, in general, higher on average for the single females. The reverse findings obtained here may be indicative of traditional discriminatory practices operating within the Irish educational system which affect the subject uptake of females at secondary and hence tertiary level.

At this point it would be of interest to contrast these results with returns based on the specification excluding occupation and industry dummies. The exclusion of this set of control variables provides a more realistic interpretation of the returns since the possession of formal educational qualifications is a pre-requisite for admission to certain occupations. As Greenhalgh (1980) points out the inclusion of occupation dummies leads to an understatement of the returns to these qualifications.

For both males and females (Tables 4.3 and 4.4, respectively) the returns to a Leaving Certificate and a university degree are more well determined than in the larger specification of Tables 4.1 and 4.2. The annual returns to a male holder of either a Leaving Certificate or a university degree is 8.5% and 42.3% respectively. The comparable female figures are 7.2% and 32.1%. In both cases these are greater in magnitude than those from the larger specification with again the female returns being on average lower than the males.

The use of educational qualifications as opposed to years in education (as some studies have employed) is based on a desire to control for the qualitative aspects of years in education. However, specifications using years in post-compulsory education as an explanatory variable in the wage equation have also been estimated. Though the full set of results is not reported here the findings are broadly consistent with those obtained from the reported specifications. The returns to an additional year in post-compulsory education for males and females, when industries and occupations are included, are 4.3% and 1.4% respectively. The comparable returns when excluding these controls are 5.2% and 2.3%.

The above discussion concerning the returns to education is contingent on the assumption that any variables omitted from the specification are orthogonal to the included ones. In terms of education and unobserved ability this may not necessarily be the case. If ability is positively correlated with education and determines wages then OLS yields an upward biased estimate of the returns to education[29]. Taubman (1976) using US survey data on twins (in an attempt to control for environmental and genetic factors) found that failing to control for ability causes a large upward bias on the educational coefficients. Hausman and Taylor (1981) exploit the use of US panel data to econometrically control for individual-specific unobservable effects which are assumed correlated with explanatory variables in their wage equation. In marked contrast to Taubman (1976) their empirical results suggest that econometric methods which control for correlation with the latent individual effects increase the schooling coefficient. On the other hand, Chowdhury and Nickell (1985) also use US panel data and extend the Hausman and Taylor (1981) econometric methodology but are unable to obtain any precise estimates for the schooling coefficients. The empirical evidence is clearly divided as to whether controlling for ability leads to an increase or a decrease in the returns to education. Since no attempt is made in this study to

---

[29] This follows directly from the formula for omitted variables outlined in Yule and Kendall (1950).

control for unobservables the cross-sectional estimates on returns to education reported here should be interpreted in a cautionary context.

Returning now to the remaining results of Tables 4.1 and 4.2. The effects of the vocational training qualifications are also recorded in tables 4.1 and 4.2. The three examined are the Group Certificate, apprenticeship qualifications and basic training qualifications[30]. In all three cases the female coefficients are badly determined and statistically insignificant. This is in part due to the small number of females in these particular categories. In terms of the males the returns to the above are 5.3%, 6.5% and 7.1% respectively.

Variables controlling for firm size were also allowed to enter the analysis. It might be argued that this variable acts as a proxy for unionisation. The larger the firm size the more likely is the possibility of union influence in regard to, for example, pay[31]. The reference group in terms of this set of dummies is firms with less than fifteen individuals. Thus, the coefficients of these dummies should be interpreted in relation to this reference group. It's evident from the results that being employed in a firm with more than one hundred employees has a large wage effect relative to a firm with less than fifteen. The effect is of the order of 25.0% for both males and females. One surprising feature of this set of results is the fact that the wage effects for males working in the medium sized category of firm is less than the effects associated with working in the smaller sized category. This finding is not however repeated for females where the effects are found to increase directly with firm size.

Promotion on-the-job is also deemed as having a positive and statistically significant effect on wages. In terms of their magnitude the effects are not much different across the sexes recording returns of 7.7% and 7.0% for the males and females respectively.

The inclusion of occupation, industry and regional dummies facilitates a ranking of these categories in terms of their effects on wages. The top occupational category for males is salaried

---

[30] The basic training is the initial training which provides an individual with the means to exercise a particular trade or profession. This might include, for example, basic training programmes which enable an individual to become an electrician, carpenter, computer programmer etc. Thus, the variable used here records any qualifications received for this particular purpose.

[31] It could, of course, be the case that large firms pay higher wages in order to discourage unionisation.

employees. The differential for this category relative to the unskilled reference group is of the order of 12.3%. The salaried employees category also ranks top for the females. The estimated differential between this group and the unskilled reference group is of the order of 42.8%. Skilled manual workers are anchored near the bottom of both the male and female occupational rankings. This may be viewed in terms of the firm specific human capital theory alluded to above. It might be argued that skilled workers in the youth labour market are in receipt of a relatively low wage in order that firms can recoup their cost outlays associated with training and this may explain the poor ranking for both sexes.

The industry rankings indicate that males in public administration are better off than in any other industry. The differential here relative to the distributive trades reference group is of the order of 11.1%. The top-ranking industry for females is the banking and insurance category. Industries with a large concentration of new foreign firms, *e.g.* those in chemical type industries, feature prominently in the top half of the industry rankings supporting the notion that jobs in these industries are relatively well paid for both males and females[32].

Reference to the coefficients on the urban and regional dummies suggest some dramatic regional differences in wages. Residing in an urban area with one thousand or more individuals is recorded as having a statistically significant effect for both males and females, 10.2% and 6.2% respectively. Furthermore, being resident in Dublin county relative to the North-West is worth a differential of nearly 13.5% for the males and 10.4% for the females with the most disadvantaged region in wage terms being the reference group itself, the North-West. This finding reflects the more favourable labour market conditions obtaining in Dublin relative to the rest of the country and could be attributed to the large number of public administration employees in this area. This finding is resonant of one of the Walsh and Whelan (1976) findings.

In general, the remaining results that feature in Tables 4.3 and 4.4 are broadly in agreement with those obtained from the larger specification and require no further comment.

---

[32] The industry and occupation dummies are statistically tested on the basis of F-tests for both the male and female equations. The F-test associated with the male equation is 1.989 ** and that of the female equation 4.338 *** (where *** denotes 1% and ** denotes 5% significance).

Table 4.1.
Male Wage Equation Estimates

Variable	Coefficient	OLS s.e.	White s.e
Constant	0.0219	0.0977	0.1138
Experience (5 or less yrs.)	0.0748***	0.0132	0.0158
Experience (more than 5 yrs.)	0.0125	0.0186	0.0131
<u>Educational Qualifications</u>			
Intermediate Certificate	-0.0406	0.0438	0.0406
Leaving Certificate	0.0424	0.0508	0.0475
University Degree	0.2899**	0.1160	0.1425
<u>Other Qualifications</u>			
Group Certificate	0.0517	0.0489	0.0554
Apprenticeship	0.0631*	0.0400	0.0370
Basic Training Qualification	0.0688*	0.0364	0.0353
<u>Job Characteristics</u>			
$15 \leq$ Firm < 50	0.1664***	0.0479	0.0534
$50 \leq$ Firm < 100	0.0968*	0.0587	0.0528
Firm ≥ 100	0.2260***	0.0422	0.0474
Promotion on Job	0.0743***	0.0323	0.0320
<u>Occupations</u>			
Professional	0.0531	0.1081	0.1094
Self-employed	-0.0697	0.1110	0.1987
Salaried employees	0.1157	0.1386	0.1363
Intermediate non-manual	-0.1867	0.0847	0.1179
Other non-manual	-0.2174*	0.0961	0.1112
Skilled manual	-0.1509	0.0778	0.1035
Semi-skilled manual	-0.1188	0.0950	0.1089
<u>Industries</u>			
Building & Engineering	0.0753	0.0538	0.0491
Transport & Communication	0.0927	0.0705	0.0615
Banking & Insurance	-0.0561	0.0805	0.0555
Public Admin. etc.	0.1055	0.0648	0.0803
Metal Manufacturing	0.0492	0.0568	0.0575
Other Manufacturing	-0.0221	0.0540	0.0563
Extractive & Chemicals	0.0886	0.0726	0.0648
<u>Regions</u>			
Dublin County	0.1270***	0.0573	0.0490
Southern Counties	0.1112**	0.0532	0.0521
Midlands Counties	0.0319	0.0571	0.0624
Leinster(excl. Dublin)	0.1334*	0.0681	0.0686
Urban	0.0968***	0.0394	0.0362
$\bar{R}^2$	0.312		
Standard Error	0.319		
Number of Cases	449		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

Table 4.2.

Female Wage Equation Estimates

Variable	Coefficient	OLS s.e.	White s.e.
Constant	-0.0611	0.1884	0.0738
Experience (5 or less yrs.)	0.0631***	0.0078	0.0077
Experience (more than 5 yrs.)	0.0247	0.0176	0.0165
Educational Qualifications			
Intermediate Certificate	-0.0653	0.0419	0.0422
Leaving Certificate	-0.0194	0.0417	0.0393
University Degree	0.2218	0.1194	0.2261
Other Qualifications			
Group Certificate	-0.0524	0.0837	0.1143
Apprenticeship	-0.0111	0.0521	0.0576
Basic Training Qualification	0.0017	0.0244	0.0244
Job Characteristics			
15 ≤ Firm < 50	0.1385***	0.0348	0.0367
50 ≤ Firm < 100	0.1491***	0.0428	0.0449
Firm ≥ 100	0.2222***	0.0293	0.0302
Promotion on Job	0.0675***	0.0220	0.0244
Occupations			
Professional	0.0640	0.1896	0.0765
Self-employed	0.1227	0.2018	0.1495
Salaried employees	0.3567*	0.2351	0.2027
Intermediate non-manual	0.0534	0.1836	0.0594
Other non-manual	-0.1122	0.1863	0.0815
Skilled manual	-0.0590	0.1880	0.0664
Semi-skilled manual	-0.0688	0.1830	0.0531
Industries			
Building & Engineering	-0.0283	0.0914	0.0888
Transport & Communication	0.0682	0.0586	0.0487
Banking & Insurance	0.1594***	0.0349	0.0350
Public Admin. etc.	0.1090***	0.0338	0.0343
Metal Manufacturing	0.1213***	0.0501	0.0465
Other Manufacturing	0.0289	0.0431	0.0400
Extractive & Chemicals	0.1382***	0.0648	0.0436
Regions			
Dublin County	0.0992***	0.0397	0.0354
Southern Counties	0.0647*	0.0380	0.0350
Midlands Counties	0.0498	0.0459	0.0392
Leinster(excl. Dublin)	0.0051	0.0551	0.0547
Urban	0.0600**	0.0282	0.0270
$\bar{R}^2$	0.391		
Standard Error	0.247		
Number of Cases	573		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

Table 4.3.

Male Wage Equation Estimates (excl. Occupations & Industries)

Variable	Coefficient	OLS s.e.	White s.e.
Constant	-0.1059	0.0650	0.6900
Experience (5 or less yrs.)	0.0808***	0.0131	0.0157
Experience (more than 5 yrs.)	0.0080	0.0188	0.0132
<u>Educational Qualifications</u>			
Intermediate Certificate	-0.0287	0.0439	0.0426
Leaving Certificate	0.0815*	0.0475	0.0486
University Degree	0.3528***	0.116	0.1335
<u>Other Qualifications</u>			
Group Certificate	0.0518	0.0491	0.0562
Apprenticeship	0.0465	0.0389	0.0349
Basic Training Qualification	0.0709**	0.0355	0.0327
<u>Job Characteristics</u>			
$15 \leq$ Firm < 50	0.1485***	0.0474	0.0520
$50 \leq$ Firm < 100	0.1063**	0.0568	0.0523
Firm ≥ 100	0.2289***	0.0377	0.0395
Promotion on Job	0.0619**	0.0319	0.0298
<u>Regions</u>			
Dublin County	0.1239**	0.0566	0.0515
Southern Counties	0.1236**	0.0526	0.0542
Midlands Counties	0.0505	0.0574	0.0638
Leinster(excl. Dublin Co.)	0.1287*	0.0687	0.0692
Urban	0.0954***	0.0396	0.0357
$\bar{R}^2$	0.290		
Standard Error	0.324		
Number of Cases	449		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

Table 4.4.

Female Wage Equation Estimates (excl. Occupations and Industries)

Variable	Coefficient	OLS s.e.	White s.e.
Constant	-0.0544	0.0575	0.0605
Experience (5 or less yrs.)	0.0643***	0.0079	0.0076
Experience (more than 5 yrs.)	0.0298*	0.0178	0.0173
Educational Qualifications			
Intermediate Certificate	-0.0236	0.0420	0.0429
Leaving Certificate	0.0697*	0.0390	0.0396
University Degree	0.2784	0.1212	0.2369
Other Qualifications			
Group Certificate	-0.0279	0.0861	0.1380
Apprenticeship	-0.0544	0.0513	0.0607
Basic Training Qualification	0.0293	0.0242	0.2406
Job Characteristics			
$15 \leq$ Firm < 50	0.1243***	0.0351	0.0372
$50 \leq$ Firm < 100	0.1364***	0.0430	0.0447
Firm ≥ 100	0.2300***	0.0281	0.0284
Promotion on Job	0.0736***	0.0226	0.0214
Regions			
Dublin County	0.1287***	0.0405	0.0363
Southern Counties	0.0677*	0.0392	0.0381
Midlands Counties	0.0477	0.0474	0.0428
Leinster(excl. Dublin)	0.0114	0.0567	0.0575
Urban	0.0599**	0.0290	0.0277
$\bar{R}^2$	0.340		
Standard Error	0.256		
Number of Cases	573		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

4.6 Diagnostic Tests

The existence of specification error in any of the above equations has serious consequences for the coefficient estimates and hence the discrimination coefficients as estimated below. This would be particularly so if the misspecification were due to the omission of relevant variables. The prime example in the literature of such an omitted variable is the motivation for work or ability which does not explicitly enter the estimating equations.

In order to have confidence in the discrimination estimates it is important to test the underlying equations. The purpose of this section is to present an array of diagnostics that explicitly test the underlying assumptions of the estimated models. These assumptions are the standard ones of the classical linear regression model as referred to in section 4.4. The error terms are assumed to have a zero mean, a constant variance and are normally distributed. The statistical tests employed here provide tests of these assumptions.

Tables 4.5 and 4.6 report the diagnostics' results for the male and female equations respectively. For each sex four equations are estimated and compared in terms of these diagnostics. In these tables (A) and (C) refer to equation (4.11) with and without the occupation and industry dummies respectively. The results of these estimated models have been the subject of extensive discussion in section 4.5. (B) and (D) refer to an alternative version of equation (4.11) that uses dummy variables for the different years of labour force experience instead of splines. The former includes occupation and industry dummies and the latter excludes them[33].

The first diagnostic to be examined is the J-test developed by Davidson and MacKinnon (1981) and suggested by the authors as a specification test. It is a non-nested test procedure and is interpreted in the context of augmenting the conditional mean (see Pagan (1984)). The proxy variable used in this case is based on predictions from the alternative or competing model. The validity of the null hypothesis (H_o) is evaluated in terms of the t-statistic associated with the proxy variable. If H_o is true, the coefficient on the proxy variable will not be significantly different from zero. This can be tested by a simple t-statistic. As Davidson and MacKinnon (1981) point out the t-statistic which is valid for testing the truth of H_o is not valid for testing the truth of H_1. Thus the roles of H_o and H_1 are reversed and the test is carried out again. Thus it is possible that both hypotheses may be rejected or both not rejected or one not rejected over the other. The J-tests recorded in tables 4.5 and 4.6 are the t-statistics associated with the proxy variables generated from the competing model. The null hypothesis alternates between the treatment of the labour force experience variable in terms of splines or in terms of dummy variables. The results provide unambiguous support in favour of the specification using splines for the male equations. The female results are, however, more ambiguous. In general, however, one would expect the continuous spline variables to dominate the discrete dummy variables and the J-test results sim-

[33] Though the coefficient estimates of these alternative models are not reported in the text they are available on request.

ply confirm this.

The second specification diagnostic is the RESET test due to Ramsey (1969). As Pagan (1984) points out, it again can be interpreted in terms of augmenting the conditional mean with a proxy variable (or proxy variables) assumed to be closely correlated with the omitted part of the conditional mean. The proxy variables, in this case, are based on the predicted values of the dependent variable from the original specification raised to a number of arbitrary powers. Simulation studies suggest that the optimal number of powers is four. The RESET test then reduces to an F-test of the significance from zero of the proxy variables. The success of the test is strongly dependent on the closeness of the correlation between these proxy constructs and the omitted part of the conditional mean[34].

In terms of the male equations the null hypothesis that there is no omitted part of the conditional mean cannot be rejected for two of the four estimated versions at the 1% level of significance. For the female equation the same null hypothesis is decisively rejected for all the estimated versions suggesting some form of potential mis-specification in the female equation.

The results of the RESET tests again establish the necessity for exercising extreme caution in the interpretation of discrimination estimates. It is obvious that regardless of which wage structure is assumed in the absence of sex discrimination the potential bias in the female coefficients will impinge upon the discrimination estimates. It should also be pointed out, in the light of the discussion in section 4.5, that the estimated returns to education and work experience for the female equations may also be rendered somewhat dubious.

The third diagnostic focuses on the assumption of constant variance[35]. The heteroscedasti-

---

[34] The test was originally described in terms of BLUS residuals by Ramsey (1969) but it was subsequently shown by Ramsey and Schmidt (1976) that this was equivalent to carrying out the above F-test in terms of the OLS residuals.

[35] It might be argued that the appropriate test statistic to use in this context is one derived from White's (1982, a) information matrix test. This information matrix test provides a test of the information matrix identity. Since LR tests, LM tests and Wald tests (which are related to F-tests and t-tests) are derived under the assumption that this identity holds, White (1982, a) suggests that the information matrix test should be used as a preliminary test to inference. Hall (1987) provides a decomposition of the information matrix test into the sum of three independent χ^2 variates. The first of which is White's (1980) direct test for heteroscedasticity with the remaining two components independent test statistics for skewness and kurtosis. However, the computation of the White heteroscedasticity test involves, for the largest specification reported here (that of table 4.10), the estimation of an auxiliary regression with a prohibitive number of explanatory variables. In view of this the more easily computable Breusch-Pagan LM test is instead used.

city test employed here is the Lagrangean Multiplier (LM) statistic developed by Breusch and Pagan (1979). This again may be interpreted in the context of augmenting a conditional moment of the original specification. In this case the conditional variance. The potential heteroscedasticity to be tested is of the form :

$$\sigma_i^2 = G_i \alpha \tag{4.12}$$

where

σ_i^2 = the variance and

G_i = is a matrix of variables, the first column of which is ones. In the context of this study the G matrix is simply composed of all the variables from the original specification.

The test statistic is calculated by first obtaining the OLS residuals from the original equation. The squared residuals from this equation are then deflated by the Maximum Likelihood estimate of the error variance from the original regression. This newly constructed variable is then regressed on the G variables and the actual test statistic is half the explained sum of squares from this regression. The resultant LM statistic is a χ^2 variate with $k - 1$ degrees of freedom where k is the number of parameters from the original equation.

The results of the Breusch-Pagan tests indicate a general rejection of the null hypothesis of homoscedasticity for most of the estimated equations. Since the standard errors employed in the analysis are the White adjusted standard errors this need not present a problem in the context of inference. The decisive rejection may be related more to mis-specification. However, due to the limitations of the data set employed obtaining a handle on a more appropriately specified model has proved almost impossible. The models estimated represent the best description of the data given this set of limitations.

A further test also aimed at detecting mis-specification has been proposed by Kiefer and Salmon (1983). Like the RESET test its advantage lies in its ease of computation but unlike the RESET test it focuses on departures from the normality assumption. This specification test is based on an Edgeworth expansion and is written as :

$$S = \frac{N}{6}(\hat{\mu}_3 - 3\hat{\mu}_1)^2 + \frac{N}{24}(\hat{\mu}_4 - 6\hat{\mu}_2 + 3)^2 \tag{4.13}$$

where N is the number of observations and the $\hat{\mu}_i$'s are the estimated sample moments based on the residuals. Following Davidson and MacKinnon (1985) the estimated sample moments are standardised by the Maximum Likelihood estimate of the standard error from the original regression. The test statistic is composed of the sum of two asymptotically independent χ^2 variates each with one degree of freedom. The test statistic has the advantage that attention can exclusively focus on either the third moment (the first part of (4.13)) or the fourth moment (the second part of (4.13)) or both. Thus one can examine the null hypothesis of skewness which has a χ^2_1 distribution independently of the null hypothesis of kurtosis (which also has a χ^2_1 distribution).

In terms of the Kiefer-Salmon normality test the female equations perform better than the male equations. Skewness is upheld for all the estimated female equations but rejected for two of the four male equations. The imposition of the restrictions on the occupation and industry dummies leads to an increase in the skewness and kurtosis test values. This is as one would expect if these restrictions are invalidly imposed. The joint test of normality is, however, clearly rejected for all equations.

The violation of the normality assumption has clear implications for the reliability of the classical statistical tests (t-tests and F-tests) which are based on the normal distribution. The coefficients, however, remain unbiased. Accepting this, however, the violation of the kurtosis assumption may be again reflecting some underlying mis-specification. In terms of the normal distribution the fourth moment is equal to three times the square of the second moment (the variance). The decisive rejection of kurtosis may thus be suggestive of some underlying heteroscedasticity again supporting the findings of the Breusch-Pagan test.

The performance of the estimated equations in terms of the diagnostics is clearly mixed. In view of this, caution should be exercised in the interpretation and use of the parameter estimates. This cautionary note should be borne particularly in mind in the context of the following section.

Table 4.5.

Diagnostics for Male Equations.

Test	(A)	(B)	(C)	(D)
J-test	1.354	2.355 *	1.535	2.827 **
RESET	2.765 *	3.377 **	1.396	1.203
Breusch-Pagan	171.053 **	187.555 **	95.101	113.132
Skewness	1.721	2.093	5.412 *	5.695 *
Kurtosis	788.880 **	686.165 **	862.732 **	783.078 **

(A) and (C) refer to the estimated equations, using splines in experience, including and excluding occupation and industry dummies respectively. (B) and (D) refer to the estimated equations, using dummies in experience, including and excluding occupation and industry dummies respectively. The non-nested J-test value is interpreted as a t-statistic. The values for **RESET** mis-specification test is interpreted as an F-test with three and $N-k$ degrees of freedom. The **Breusch-Pagan** test for heteroscedasticity is interpreted as a χ^2 variate with degrees of freedom equal to the number of parameters estimated in the original equations less one. The normality tests of skewness and kurtosis are based on the Keifer-Salmon test statistic and are interpreted as χ^2 variates with one degree of freedom each. ** and * denotes significance at the 1% and 5% level of significance respectively.

Table 4.6.

Diagnostics for Female Equations.

Test	(A)	(B)	(C)	(D)
J-test	1.429	1.483	2.304 *	1.041
RESET	8.698 **	8.084 **	5.582 **	4.760 **
Breusch-Pagan	149.649 *	156.552 *	112.673	112.528
Skewness	0.004	0.001	0.553	0.853
Kurtosis	88.426 **	78.011 **	169.106 **	142.407 **

See Table 4.5 for a full description of these tests. ** and * denotes significance at the 1% and 5% level of significance respectively.

4.7 Discrimination Estimates

Section 4.2 outlined in detail the methodology to be adopted in estimating the wage effects of sex discrimination. A male wage structure is assumed to best characterise the conditions that obtain in the absence of gender wage discrimination. Estimates of the explained and unexplained portions of the differential are calculated for the models including occupation and industry controls and for those excluding these controls. Standard errors are also calculated for both parts of the observed gender differential.

The observed gender wage differential in logarithms is 0.1037 at the mean suggesting that young males earn on average 10.9% more than young females. The unexplained part due to differing characteristics ($\bar{X}'\Delta\hat{\beta}$) is 0.0305. This constitutes a little under 30% of the observed differential. Thus the greater proportion of the wage differential is explained by the possession of differing characteristics with a relatively small amount due to the presence of unexplained factors which is taken, in terms of this analysis, to be wage discrimination. The asymptotic standard error associated with 0.0305 is 0.0389 suggesting that the unexplained differential is not statistically significant. That part of the observed differential explained by characteristics ($\Delta\bar{X}\hat{\beta}^{m}$) is 0.0732 and its associated standard error is 0.0300. The explained differential is clearly different from zero at the 5% level of statistical significance.

Examining the average may provide a misleading picture of how the differential behaves across different types or categories of individuals. Following Stewart (1983) table 4.7 contains differential calculations based on a male wage structure for different stylised female workers. The first such stylised individual falls into the base group (*i.e.* has an amount of labour force experience equal to the female mean in the sample and scores zero on all the binary variables controlled for in the analysis). The differential due to discrimination in this case is estimated at 12.9% but again is not recorded as being statistically significant. The remainder of the table reports deviations from this base set of characteristics. Each deviation is examined by itself alone with the objective of trying to establish whether there exists a large variation in the differential across differing characteristics. Asymptotic standard errors are also recorded to establish statistical significance. Though the variation is found to be large in regard to some characteristics all but one is found to be statistically insignificant. The differential is seen to decline with labour force experience. The more experience an individual possesses the smaller is the unexplained differential. Again, however, this effect is not found to be statistically significant.

The only significant (and the largest) unexplained differential is obtained for the Leinster region (excl. Dublin Co.). The differential is of the order of 28% and its magnitude is explained by the vast difference between the male and female coefficients associated with this regional dummy. This large differential could be explained by the fact that males resident in the Leinster region are more able to commute to well paying jobs in Dublin county than are females residing in this region. This, however, can only be offered as a tentative explanation for what is a rather odd result.

Another interesting feature of table 4.7 is that there exists some evidence of "reverse" discrimination in terms of three occupational categories; the self-employed, salaried employees and the intermediate non-manual categories. In terms of the self-employed discrimination usually takes the form of consumer motivated discrimination. An interpretation for the result recorded here is that consumers are more likely to discriminate against young self-employed male workers than their female counterparts. The magnitude of the effect in this case is a little under 7% but is not statistically significant.

The negative effect in the intermediate non-manual category (an effect of over 11% in favour of the females) helps explain the relatively low average value recorded in the first row of table 4.7. Since the mean discrimination coefficient, in this case, is the difference in coefficients weighted by the female mean characteristics, the mean estimate will clearly be influenced by the proportions in certain categories. Since over 70% of all the females in the sample are in the intermediate non-manual category this clearly has a dampening effect on the mean estimate. This clearly highlights one of the dangers associated with using this approach.

Table 4.8 contains calculations of the gender wage differential based on the estimated models of tables 4.3 and 4.4 (*i.e.* those excluding the occupation and industry dummies). As anticipated in section 4.2 the mean estimate of the unexplained differential increases dramatically and becomes statistically significant. The $\ln(1 + D)$ estimate is 0.0847 and with a standard error of 0.0264 is statistically different from zero at the 1% level of significance. The portion of the differential explained by characteristics falls and becomes statistically insignificant. This clearly illustrates the role played by the occupation and industry compensating differentials. Failure to control for these effects clearly distorts the wage discrimination estimates. In view of this, the more realistic estimates of discrimination are assumed obtained from the larger occupation/industry specification[36].

As in table 4.7 base calculations and deviations from the base are also reported. Again, despite the significance of the mean differential, none of the remaining set of calculations is statistically significant. The findings of table 4.7 are, more or less, repeated here. However, in contrast, it appears evident that the larger the firm size the less likely is there of a wage differential in favour of males. Furthermore, the Leinster region differential though still relatively large becomes statistically insignificant.

[36] Bearing in mind the important caveat in regard to occupational endogeneity.

Table 4.7.

Gender Differential Estimates

Characteristic	*ln*(1 + D)	1 + D	ASE of *ln*(1 + D)
Mean	0.0305	1.0310	0.0389
Base	0.1212	1.1288	0.1252
Educational Qualifications			
Intermediate Certificate	0.1459	1.1571	0.1267
Leaving Certificate	0.1829	1.2007	0.1317
University Degree	0.1892	1.2083	0.3004
Other Qualifications			
Group Certificate	0.2253	1.2527	0.1708
Apprenticeship	0.1954	1.2158	0.1477
Basic Training Qual.	0.1882	1.2071	0.1358
Job Characteristics			
$15 \leq$ Firm < 50	0.1490	1.11607	0.1291
$50 \leq$ Firm < 100	0.0689	1.0713	0.1351
Firm ≥ 100	0.1250	1.1331	0.1381
Promotion on Job	0.1280	1.1365	0.1349
Occupations			
Professional	0.1103	1.1166	0.1136
Self-employed	-0.0713	0.9312	0.2158
Salaried employees	-0.1198	0.8871	0.2309
Intermediate non-manual	-0.1190	0.8878	0.1065
Other non-manual	0.0159	1.0160	0.1066
Skilled manual	0.0292	1.0296	0.1080
Semi-skilled manual	0.0711	1.0737	0.1098
Industries			
Building & Engineering	0.2248	1.2521	0.1469
Transport & Communication	0.1341	1.1567	0.1341
Banking & Insurance	-0.0943	0.9100	0.1316
Public Admin. etc.	0.1176	1.1248	0.1280
Metal Manufacturing	0.0491	1.0503	0.1266
Other Manufacturing	0.0702	1.0727	0.1292
Extractive & Chemicals	0.0716	1.0742	0.1286
Regions			
Dublin County	0.1489	1.1606	0.1277
Southern Counties	0.1677	1.1826	0.1204
Midlands Counties	0.1033	1.1088	0.1270
Leinster(excl. Dublin)	0.2495*	1.2834	0.1296
Urban	0.1579	1.1710	0.1239
Work Experience			
Five years	0.1418	1.1523	0.1300
Six years	0.1296	1.1384	0.1291
Seven years	0.1174	1.1246	0.1315
Eight years	0.1053	1.1110	0.1372

See Table 4.8 for a full explanation. *** denotes significance at 1% level, ** significance at the 5% level and * significance at the 10% level.

Table 4.8.

Gender Differential Estimates (excl. Occupations and Industries)

Characteristic	*ln* (1 + D)	1 + D	ASE of *ln* (1 + D)
Mean	0.0847***	1.0884	0.0264
Base	0.0016	1.0016	0.0800
Educational Qualifications			
Intermediate Certificate	-0.0035	0.9965	0.0781
Leaving Certificate	0.0133	1.0134	0.0806
University Degree	0.0759	1.0788	0.2768
Other Qualifications			
Group Certificate	0.0813	1.0847	0.1513
Apprenticeship	0.1025	1.1079	0.0916
Basic Training Qual.	0.0432	1.0441	0.0860
Job Characteristics			
15 ≤ Firm < 50	0.0257	1.0260	0.0806
50 ≤ Firm < 100	-0.0285	0.9719	0.0843
Firm ≥ 100	0.0004	1.0004	0.0781
Promotion on Job	-0.0101	0.9899	0.0787
Regions			
Dublin County	-0.0033	0.9967	0.0854
Southern Counties	0.0574	1.0591	0.0872
Midlands Counties	0.0044	1.0044	0.0949
Leinster(excl. Dublin)	0.1188	1.1261	0.0995
Urban	0.0370	1.0377	0.0781
Work Experience			
Five years	0.0301	1.0306	0.0889
Six years	0.0083	1.0083	0.0854
Seven years	-0.0136	0.9865	0.0872
Eight years	-0.0354	0.9652	0.0943

The first row records the mean differential estimate. The second row records the differential estimate for an individual with a base set of characteristics. The third and subsequent rows record deviations from the base characteristics and are allowed to occur singly. The fourth column in the above table reports the asymptotic standard errors (ASE) of the estimated differentials. * denotes significance at the 1% level, ** significance at the 5% level and * significance at the 10% level.**

It might be asked how these estimates compare with the UK evidence. Since Greenhalgh (1980) employed a roughly similar methodology[37] to the one adopted in this chapter the estimates contained therein will be used for comparison purposes. The data set used was the 1975 GHS and the closest comparable group Greenhalgh examined of interest in terms of this study is the under thirty single men/single women subset. In the context of this group an unexplained residual estimate of 10% was obtained. However, no standard errors were calculated for this or any other of the estimates recorded there. Since this is a 1975 estimate and is based on a group with an older terminal age than the one employed here a cross country comparison between the two must remain at least tentative and at most crude. It must remain a matter of conjecture as to how much of the seven percentage points that separate the two estimates could be explained by (i) the passing of time and the influence of equal pay legislation, (ii) differences in the structure of the Irish and UK labour markets and (iii) the fact that the specification estimated here controls explicitly for certain job characteristics (*e.g.* firm size and on-the-job promotion) in a way that the Greenhalgh study did not. One may speculate but it would be a major surprise if more recent UK estimates didn't more closely mirror the Irish estimates.

In general tables 4.7 and 4.8 reveal little statistical difference between the coefficients of the male and female equations. In the light of this a more parsimonious model using the pooled sample of males and females with a sex dummy is estimated. The sex dummy adopts a value of 1, if male and 0 otherwise[38]. The full sample specification is estimated with occupation and industry controls and the results of this exercise are contained in Table 4.9. Most of the coefficient estimates are in line with those reported in tables 4.1 and 4.2. The coefficient of most interest in table 4.9 is the sex coefficient. This suggests that males on average earn 7.5% more than females. Of the full set of interactions attempted only two are statistically significant at conventional lev-

---

[37] The Greenhalgh methodology as pointed out in chapter two did not assume a male wage distribution in the absence of discrimination but took the geometric mean of both male and female coefficient estimates to calculate a discrimination estimate.

[38] Chow tests are carried out to statistically test whether the two separate sex models fit the data better than the pooled sample model constrained by the inclusion of a sex dummy. Chow tests for the specifications including occupation and industry dummies and excluding these controls are calculated. The null hypotheses are the sex dummy constrained models. The resultant F-tests are 1.046 ($F(31,958)$) and 0.645 ($F(17,986)$). Neither of the null hypotheses can be rejected.

els of significance. As table 4.10 indicates the two significant interactive terms are sex and the Banking and Insurance industry category and sex and the intermediate non-manual occupational category. The coefficients of the interactive terms are interpreted as the *ceteris paribus* differences between the male and female coefficients for the given categories. For both the categories in question the female coefficients are statistically and significantly larger than the male coefficients. In terms of table 4.7 above a similar finding is reported for these two categories with the effect, however, statistically insignificant. The difference in statistical significance is due to the fact that the results of table 4.7 explicitly assume a male wage structure. For completeness, diagnostics for these two equations are reported in table 4.11.

Table 4.9.

Pooled Wage Equation Estimates

Variable	Coefficient	OLS s.e.	White s.e.
Constant	0.0022	0.0762	0.1041
Sex	0.0722**	0.0257	0.0282
Experience (5 or less yrs.)	0.0651***	0.0069	0.0074
Experience (more than 5 yrs.)	0.0217**	0.0123	0.0099
Educational Qualifications			
Intermediate Certificate	-0.0316	0.0293	0.0304
Leaving Certificate	0.0284	0.0307	0.0312
University Degree	0.3049***	0.0798	0.1166
Other Qualifications			
Group Certificate	0.0471	0.0372	0.0465
Apprenticeship	0.0560*	0.0278	0.0292
Basic Training Qualification	0.0424**	0.0197	0.0196
Job Characteristics			
$15 \leq$ Firm < 50	0.1465***	0.0284	0.0322
$50 \leq$ Firm < 100	0.1255***	0.0347	0.0344
Firm ≥ 100	0.2251***	0.0242	0.0263
Promotion on Job	0.0682***	0.0185	0.0183
Occupations			
Professional	-0.0636	0.0745	0.1034
Self-employed	-0.0472	0.0823	0.1644
Salaried employees	0.0997	0.1044	0.1353
Intermediate non-manual	-0.1120	0.0654	0.1035
Other non-manual	-0.2507**	0.0699	0.1034
Skilled manual	-0.1420	0.0636	0.0953
Semi-skilled manual	-0.1520	0.0704	0.1004
Industries			
Building & Engineering	0.0776*	0.0407	0.0405
Transport & Communication	0.0792**	0.0445	0.0378
Banking & Insurance	0.1224***	0.0326	0.0312
Public Admin. etc.	0.0894***	0.0304	0.0329
Metal Manufacturing	0.0683*	0.0364	0.0365
Other Manufacturing	-0.0032	0.0334	0.0342
Extractive & Chemicals	0.1199***	0.0469	0.0397
Regions			
Dublin County	0.1174***	0.0329	0.0299
Southern Counties	0.0893***	0.0311	0.0312
Midlands Counties	0.0463	0.0356	0.0371
Leinster(excl. Dublin)	0.0805*	0.0427	0.0457
Urban	0.0762***	0.0231	0.0220
$\bar{R}^2$	0.357		
Standard Error	0.281		
Number of Cases	1022		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

Table 4.10.
Pooled Wage Equation Estimates with Interactions

Variable	Coefficient	OLS s.e.	White s.e.
Constant	-0.0646	0.0784	0.1018
Sex	0.1434***	0.0330	0.0333
Experience (5 or less yrs.)	0.0660***	0.0069	0.0074
Experience (more than 5 yrs.)	0.0215**	0.0122	0.0099
Educational Qualifications			
Intermediate Certificate	-0.0375	0.0292	0.0299
Leaving Certificate	0.0220	0.0307	0.0303
University Degree	0.2831**	0.0796	0.1165
Other Qualifications			
Group Certificate	0.0411	0.0371	0.0465
Apprenticeship	0.0619**	0.0279	0.0298
Basic Training Qualification	0.0320	0.0198	0.0199
Job Characteristics			
15 ≤ Firm < 50	0.1480***	0.0283	0.0318
50 ≤ Firm < 100	0.1244***	0.0346	0.0339
Firm ≥ 100	0.2264***	0.0241	0.0262
Promotion on Job	0.0705***	0.01838	0.0181
Occupations			
Professional	-0.0023	0.0759	0.1003
Self-employed	-0.0181	0.0825	0.1599
Salaried employees	0.1559	0.1051	0.1292
Intermediate non-manual	-0.0368	0.0701	0.0988
Other non-manual	-0.2082**	0.0706	0.1005
Skilled manual	-0.1395	0.0633	0.0945
Semi-skilled manual	-0.1208	0.0707	0.0980
Industries			
Building & Engineering	0.0704*	0.0405	0.0405
Transport & Communication	0.0781**	0.0442	0.0374
Banking & Insurance	0.1451***	0.0351	0.0335
Public Admin. etc.	0.0945***	0.0303	0.0329
Metal Manufacturing	0.0714**	0.0362	0.0361
Other Manufacturing	-0.0017	0.0332	0.0340
Extractive & Chemicals	0.1166***	0.0467	0.0392
Regions			
Dublin County	0.1207***	0.0328	0.0295
Southern Counties	0.0939***	0.0310	0.0307
Midlands Counties	0.0460	0.0354	0.0368
Leinster(excl. Dublin)	0.0806*	0.0425	0.0450
Urban	0.0783***	0.0230	0.0220
Interaction Terms			
Sex×Banking & Ins.	-0.1688**	0.0573	0.0685
Sex×Int. non-manual	-0.1361***	0.0501	0.0525
$\bar{R}^2$	0.364		
Standard Error	0.279		
Number of Cases	1022		

Statistical inference is based on the White standard errors. Two-tailed test of significance are employed and * denotes significance at the 1% level, ** at the 5% level and * at the 10% level.**

Table 4.11.

Diagnostics for Pooled Wage Equations

Test	(A)	(B)	(C)	(D)
J-test	0.688	2.483 **	0.315 *	2.782 **
RESET	8.740 **	9.656 **	8.022 **	8.824 **
Breusch-Pagan	301.398 **	326.377 **	319.699 **	345.164 **
Skewness	3.646	3.467	2.713	2.506
Kurtosis	1403.277 **	1402.493 **	1316.547 **	1081.530 **

(A) and (C) refer to the estimated equations using splines in experience with occupation and industry dummies. The results for which are reported in tables 4.9 and 4.10 respectively. (B) and (D) refer to the estimated equations using dummies in experience with occupation and industry dummies. The results for which are not reported here but are available on request. The non-nested **J-test** value is interpreted as a t-statistic. The values for the **RESET** mis-specification test is interpreted as an F-test with 3 and $N-k$ degrees of freedom. The **Breusch-Pagan** test for heteroscedasticity is interpreted as a χ^2 variate with degrees of freedom equal to the number of parameters estimated in the original equations less one. The normality tests of skewness and kurtosis are based on the Keifer-Salmon test statistic and are interpreted as χ^2 variates with one degree of freedom each. ** and * denotes significance at the 1% and 5% level respectively.

4.7 Conclusions

The estimation of individual level wage equations has provided an opportunity to examine the returns to both labour force experience and educational qualifications. By and large, the results are broadly compatible with the predictions of human capital theory. The returns to on-the-job training are greater for young males in the first five years of labour force experience than for young females. However, the returns are found to diminish for the males by a more dramatic amount in the subsequent years than is the case for the young females. Returns to educational qualifications are on average higher for males than females and this might be interpreted as reflecting some form of discrimination in terms of female access to certain subjects within the Irish educational system. The differences in returns to educational qualifications are not found to be statistically significant.

The general view that emerges from the above exercise is that the data do not provide any convincing evidence in support of wage discrimination in the context of the labour market for young workers. This should not be interpreted as suggesting that sex discrimination *per se* is an absent phenomenon from the Irish labour market for young workers. It can be stated that the results of the above analysis provide scant statistical evidence in support of a discrimination effect that originates through wage differences. However, the approach adopted would not be expected to detect forms of sex discrimination that occur through, for example, the existence of barriers to occupational entry or employer motivated discrimination in terms of on-the-job promotion offers. One may conjecture that the small magnitude recorded for the unexplained wage differential may be attributable to the success of equal pay legislation. Wages are the one obvious variable that can be easily regulated by anti-discrimination legislation. Introducing and implementing legislation to remove wage discrimination may be a far easier task than removing certain other forms of employer motivated discrimination that manifest themselves through, for example, promotional offers to females. Failure to detect sex discrimination in the form of wage effects does not necessarily imply the absence of sex discrimination in its other forms.

Therefore, the results obtained cannot claim to represent the definitive statement on sex discrimination in the labour market for young workers in Ireland. Since the focus of attention has been young single workers the discrimination effect measured here does not relate to any discrimination that may occur as a consequence of female labour force intermittency. Nor has the focus here been on other forms of discrimination that may arise as a consequence of either occupational segregation or unequal access to promotion. The absence of wage sex discrimination cannot be interpreted as *prima facie* evidence against the existence of any of these other types of discrimination.

In the light of these findings the subsequent chapters attempt to treat occupations in a more sensitive fashion than has been the case in this chapter. In particular the issue of occupational endogeneity is more extensively explored in the following chapter and in chapter six the wage effects of occupational segregation is assessed. The sample employed in this chapter is also expanded in order to facilitate a more worthwhile analysis of occupational wage effects. This inevitably means that certain variables for which there are too few observations must be dropped. These include the vocational training variables and the promotion variables.

Chapter Five

Gender Wage Discrimination with Endogenous Occupations

5.1 Introduction

Chapter four focused on gender wage effects treating the occupational levels of males and females as exogenous. As stated in that chapter, and also in chapter two, the literature is replete with examples of such studies where the unexplained differential between two reduced form wage equations is assumed to approximate a wage discrimination effect. One of the major limitations of such studies is that the occupational effects are controlled through intercept shifts in the wage equation. The estimation of separate occupational wage equations allowing for differing returns to characteristics across occupations represents a clear advance. This is particularly so if there is suspicion that the mean discrimination effect conceals the presence of a larger intra-occupational effect.

However, a major problem posed by the estimation of occupational wage equations relates to the possible existence of some selection process that determines the observed occupational sample. If the disturbances in the occupational wage equations are correlated with the disturbance term in the occupational selection equation then conventional estimation techniques, like OLS, provide biased and inconsistent parameter estimates. This has clear implications for the estimated discrimination effects. Methods designed to correct for such selectivity bias have been suggested in the literature and applied to the area of labour supply, migration (Robinson and Tomes, (1982)) and union endogeneity (Duncan and Leigh (1980)). Few studies have analysed the effects of selectivity bias on the discrimination estimates[39] bearing particular emphasis on the

[39] Zabalza and Arrufat (1983) and Dolton and Makepeace (1986 and 1987, a) are notable exceptions in the field of sex discrimination with Reimers (1983) providing an exception in a race discrimination study. The former three studies focus on the sample selectivity effects of labour force participation.

effects of occupational selection. Dolton, Makepeace and Van der Klaauw (1987) have examined occupational wages and the effects of sample selectivity in a polychotomous occupational framework but without explicit reference to intra-occupational wage discrimination effects.

One of the main objectives of this chapter is to explore gender and occupational wage differentials within a dichotomous non-manual/manual framework and to establish the effects, if any, of occupational sample selectivity on sex discrimination estimates. A second objective is to statistically test the proposition of occupational exogeneity.

Two contrasting econometric methods are employed to control and test for the potential endogeneity of occupational status. One is an Instrumental Variable (IV) estimator proposed by Dubin and McFadden (1984) and refined for use by Duncan and Leigh (1985) in the context of union endogeneity. The other is the widely used two-stage selectivity bias correction method based on the Mill's ratio and proposed by Heckman (1976). Testing for occupational exogeneity in the former case is effected through the calculation of a Hausman test (see Hausman (1978)). In the latter case the statistical test for endogeneity is derived from Melino (1982) who provides a Lagrange multiplier test based on the t-statistic of the predicted selectivity bias term.

The econometric issues raised by the analysis should not hide important economic policy issues. Foremost among these is the question of whether the magnitude of the unexplained gender wage differential varies markedly across manual and non-manual occupational sectors. A second question relates to the age of the sample of workers used in the analysis. Some theoretical models highlight the role played by female labour force interruption and subsequent skill depreciation in providing an explanation for female wage disadvantage.[40] In the context of young workers of single status one may be surprised to find evidence of wage based discrimination in any occupational sector. The detection of such an effect has clear and disturbing implications for the transition of young female workers into the adult labour market.

Finally, though not the major objective of this chapter the framework employed implies the existence of a structural occupational model. The estimation of this, it is hoped, will provide

[40] See Mincer and Polachek (1974) and Polachek (1981).

some evidence as to the key determinants of non-manual occupational attachment and establish the role played by the non-manual/manual wage gap.

It should, of course, be stated that the manual/non-manual split may not be the most satisfactory categorisation to employ. However, breaking the occupational groupings down into a finer classification would have implications for both the sample size in each category and the IV econometric technique proposed for use in this chapter. A finer classification is employed in chapter six.

The layout of the chapter is as follows: sections 5.2 to 5.4 outline the methodology to be employed and provide a comparison of the different econometric methods to be used. Section 5.5 deals with the data set to be used and sections 5.6 to 5.8 concentrate on the wage equation estimation and the exogeneity results. Section 5.9 compares the unexplained gender wage differentials and section 5.10 provides the structural estimates of the occupational model. Section 5.11 offers some conclusions.

5.2 Methodology

The model describing the determination of non-manual and manual occupational attachment and wages is given by the following set of equations.

$$Y_i = K_i \gamma + \varepsilon_i \tag{5.1}$$

$$w_{ni} = X_{ni} \beta_n + \eta_{ni} \tag{5.2}$$

$$w_{mi} = X_{mi} \beta_m + \eta_{mi} \tag{5.3}$$

where

$i = 1, \ldots\ldots, T$,

T = the number of individuals,

n and m subscripts refer to non-manual and manual occupational categories respectively,

$\mathbf{Y}_i$ is the latent variable for the i^{th} individual capturing the determinants of occupational attachment,

$\mathbf{w}_i$ is the natural log of the net hourly wage for i^{th} individual,

$\mathbf{X}_i$ is a vector of characteristics that determines the i^{th} individual's net hourly wage,

$\mathbf{K}_i$ is a vector of characteristics that determines the i^{th} individual's occupational attachment and

ε_i, η_{ni} and η_{mi} are error terms.

$\mathbf{Y}_i$ is an unobserved latent dependent variable in the reduced form criterion function that predicts occupational attachment. The criterion function itself is obtained by substituting the reduced form wage equations of (5.2) and (5.3) into a structural occupational attachment equation where the non-manual/manual wage differential enters as an explanatory variable. Thus, since wages are assumed to determine occupational attachment all the explanatory variables that influence the individual's wage also influence occupational attachment through a reduced form equation. Though a structural occupational equation is to be estimated equation (5.1) is the reduced form occupational attachment equation and not the structural equation of the model. In terms of the above equations all the variables in the $\mathbf{X}_i$ vectors of equations (5.2) and (5.3) are contained in the $\mathbf{K}_i$ vector of equation (5.1) which also includes additional variables from the structural occupational equation.

The dichotomous realisation of the unobserved $\mathbf{Y}_i$ is assumed to be a dummy indicator variable, $\mathbf{I}_i$. If $\mathbf{I}_i = 1$, then, the observed individual is in a non-manual job and if $\mathbf{I}_i = 0$, the worker is attached to a manual job. Invariably occupational attachment is described in terms of utility gain. One would expect, *ceteris paribus*, that if an individual's utility gain associated with non-manual work exceeded that associated with manual work the individual would select a non-manual job. However, in the presence of job rationing, barriers to occupational entry and other forms of discrimination *etc.* this need not necessarily be the case. Thus, for an individual drawn at random from the population as a whole if $\mathbf{Y}_i \geq 0$ one observes the non-manual wage and status; otherwise one observes the manual wage and status.

Equations (5.1) to (5.3) represent the model to be estimated. The empirical implementation of this model is effected by two contrasting econometric techniques. These techniques are to be described in the subsequent sub-sections.

5.3 Heckman Procedure

The wage equations of (5.2) and (5.3) cannot be validly estimated separately by OLS since estimation would be on the basis of a truncated sample. The truncation of the sample follows from the fact that the non-manual wage is unobserved for the manual worker and vice-versa. As Duncan (1983) points out if both manual and non-manual wages are observed for each individual drawn at random from the population, then the application of OLS (with the standard set of caveats) is valid. In reality such circumstances rarely, if ever, occur and in the presence of such truncation OLS is invalid. However, Heckman (1976 and 1979) provides a method for estimating in the presence of such truncation. The regression equations of (5.2) and (5.3) may be expressed as

$$E(w_{ni} \mid X_{ni}, Y_i \geq 0) = X_n \beta_n + E(\eta_{ni} \mid Y_i \geq 0) \tag{5.2*}$$

$$E(w_{mi} \mid X_{mi}, Y_i < 0) = X_m \beta_m + E(\eta_{mi} \mid Y_i < 0) \tag{5.3*}$$

where all the elements are as defined above with E( ) depicting the expectations operator. Heckman (1979) points out that the straight application of OLS to such cases as those depicted in (5.2) and (5.3) suffers from two sources of misspecification; one due to omitted variables, the other to heteroscedasticity. Heckman proposes the use of proxy constructs designed to take in to consideration the truncated nature of the error terms depicted in (5.2*) and (5.3*). The regression equations (5.2*) and (5.3*) may then be restated as:

$$E(w_{ni} \mid X_{ni}, Y_i \geq 0) = X_n \beta_n + \theta_n \lambda_{ni} \tag{5.4}$$

$$E(w_{mi} \mid X_{mi}, Y_i < 0) = X_m \beta_m + \theta_m \lambda_{mi} \tag{5.5}$$

where

$$\lambda_{ni} = \frac{\phi(K_i\gamma)}{\Phi(K_i\gamma)} \tag{5.6}$$

$$\lambda_{mi} = -\left[\frac{\phi(K_i\gamma)}{1-\Phi(K_i\gamma)}\right] \tag{5.7}$$

$\phi(\ )$ and $\Phi(\ )$ are the density and distribution functions of a standard normal variable.

Olsen (1982) highlights the necessity of imposing some form of structure on the problem of correcting for selectivity bias without which, he argues, the problem is insoluble. This raises the issue of identification of the selectivity effect. In the context of the empirical union endogeneity literature identification creates a clear problem. All variables that influence the wage also influence union attachment and identification of the selectivity effect relies on the functional form. Since, the Mill's ratio is a non-linear function of the exogenous variables in the probit equation the same set of regressors can be used in (5.1) as in (5.2) or (5.3) without encountering collinearity. However, a condition required for the identification of the selectivity effect in the two-step framework outlined is the availability of some variable that shifts the probability of observing the dependent variable without shifting the mean of the dependent variable. For the purposes of this study a set of parental background dummies are included designed to shift the probability of occupational attachment but do not enter the wage equation.[41] An alternative solution to the identification problem lies in the use of non-linearities in the exogenous variables, *e.g.* squared or interactive effects, in order to identify the relationship. Since, an investigator rarely possesses any intuition regarding the appropriate functional form Olsen (1980) dismisses this approach as relatively unappealing.

The standard approach to estimating the above model in the Heckman two-stage framework is to apply probit[42] analysis to the reduced form criterion function of (5.1) yielding estimates $\hat{\gamma}$. Insert these estimates into (5.6) and (5.7) to obtain proxy constructs designed to control for the truncated nature of the error terms in the non-manual and manual wage equations. The second

---

[41] However, studies examining the wage effects of parental background are not uncommon and some have used father's occupational status as an explanatory variable in wage equations in order to assess differences in the private rates of return to education across socio-economic groups (see Papanicolaou and Psacharopoulos (1976)).

[42] A good survey of the properties of probit models is given in either Amemiya (1981) or Maddala (1983).

stage of the procedure involves the application of OLS to the heteroscedastic regression equations of (5.4) and (5.5). All that is required to obtain consistent estimates of the parameters is that the error term in the reduced form of (5.1) is Gaussian and that the expectations (conditional on ε_i) of the error terms of the wage equations are linear (see Olsen (1980) or Duncan (1983)).

The estimated variances used for inference purposes in terms of the Heckman procedure are corrected to take into consideration the fact that the proxies used are estimated and not actual. Maddala (1983) demonstrates how ignoring the fact that γ is estimated leads to an under-estimation of the true variances. The correction required is due to Lee *et al.* (1980) and is given in the appendix to chapter eight of Maddala (1983). However, in terms of this chapter's estimates little difference was noted between the OLS variance and the true variance calculated as in Maddala (1983).

The interpretation of the coefficients associated with (5.6) and (5.7) is always difficult. As Lee (1978) points out the θ_n and the θ_m coefficient estimates can be analytically shown to represent the covariances between the reduced form error term and that of the relevant non-manual and manual wage equations. Therefore, these terms may not be amenable to an explicit economic interpretation. However, Dolton and Makepeace (1987, b) offer some insight into the economic interpretation of such effects and a wider discussion of this is avoided until the results section.

The statistical test for exogeneity in this framework is provided by Melino (1982) who shows that the Heckman test is equivalent to a Lagrange multiplier test of the null hypothesis of no sample selection bias. The LM test is derived as the square of the t-statistic on the selectivity bias term using the uncorrected OLS variance which is consistent under the null hypothesis of no sample selection bias. The resultant test statistic is a χ^2 variate with one degree of freedom.

This two-step procedure has not been free of criticism. In particular Lee (1982) suggests that the imposed normality assumption on the error term of the criterion function may have serious implications for the detection of selectivity bias. A failure to detect such bias when present may be related to a misspecification of normality in the error term. Lee (1983) suggests a selectivity bias correction method that allows for more general distributional assumptions. However, the problem of having to make some distributional assumption is not avoided. It is this particular problem which has forced attention to turn towards distribution free estimators among which is included the IV estimator of the following section.

5.4 IV Procedure

The IV estimation technique employed here follows closely that proposed by Duncan and Leigh (1985) which provides an extension of the Dubin and M^c Fadden (1984) methodology. Retaining the notation used above the full sample wage equation may be written as

$$w_i = I_i w_{ni} + (1 - I_i) w_{mi} \tag{5.8}$$

Substituting in for the non-manual and manual wages using (5.2) and (5.3) yields

$$w_i = (I_i X_{ni})\beta_n + ((1 - I_i) X_{mi})\beta_m + v_i \tag{5.9}$$

or

$$w_i = Z_{ni} \beta_n + Z_{mi} \beta_m + v_i \tag{5.10}$$

where

$$v_i = I_i \eta_{ni} + (1 - I_i)\eta_{mi}$$

and the error terms are assumed to have the following properties:

$$E(v_i) = 0, \tag{5.11}$$

$$\mathrm{var}(v_i^2) = \sigma_i^2 \tag{5.12}$$

The fully interactive model suggested by (5.9) or (5.10) allows returns to the variables to vary across occupational sectors. However, the use of OLS in estimating (5.10) is invalidated by the fact that $E((Z_{ni} : Z_{mi})v_i) \neq 0$. As Duncan and Leigh (1985) show in order to estimate (5.10) uning the IV procedure the stringent condition that the joint density functions $g(\varepsilon_i, \eta_{mi})$ and $g(\varepsilon_i, \eta_{mi})$ are equal is imposed. This implies that the error generating process that characterises the wage equations in the two sectors is approximately the same for the first two moments of the distribution. This is necessary (as the authors show in an appendix) to ensure that the properties of (5.11) and (5.12) are satisfied.

A necessary criterion for admissible instruments is high correlation with the regressor in question, *i.e.* occupation. Duncan and Leigh (1985) suggest that natural instruments to use in the IV estimation of (5.10) are the expected values of the explanatory variables, $E(Z_{mi}) = P_i X_{mi}$ and $E(Z_{mi}) = (1 - P_i)X_{mi}$ where P_i = prob(I_i = 1) with P_i calculated from the reduced form of (5.1) using a probit. This is identical to the first stage of the Heckman estimator. Instruments are then formed by interacting the predicted probabilities $\hat{P}_i$ with the actual X_{ui} and X_{mi} variables. Define the instruments calculated in this manner by the matrix **W** and denote the vector of natural log of the net hourly wage by **y**. If the matrix $(Z_{ui}:Z_{mi})$ is denoted more simply by **Z**; then the well known IV coefficient estimator is given by

$$\hat{\beta}_{IV} = (W'Z)^{-1}W'y \tag{5.13}$$

The estimator for the variance is modified to take into consideration the presence of heteroscedasticity.

$$var(\hat{\beta}_{IV}) = (W'Z)^{-1}W'\hat{\Omega}W(Z'W)^{-1} \tag{5.14}$$

where

$$\hat{\Omega} = diag((y - Z'\hat{\beta}_{IV})'(y - Z'\hat{\beta}_{IV}))$$

This is the covariance matrix estimator based on White (1982, b) and is consistent under both the null hypothesis of homoscedasticity and the alternative of heteroscedasticity regardless of its structure.

The advantage the IV approach possesses over the Heckman procedure is the fact that no distributional assumptions enter the second stage of estimation. Though a normality assumption is necessary to obtain the predicted probabilities in the IV case this assumption does not enter the wage equation estimation as with the Heckman two-step procedure[43].

The statistical test for occupational endogeneity in the IV case is provided by Hausman (1978). This statistically compares an estimator (OLS) that is consistent and efficient under the null hypothesis of exogeneity but inconsistent under the alternative hypothesis of endogeneity against an alternative estimator (IV) that is consistent under both the null and the alternative. However the IV estimator may be inefficient if the correlations between instruments and the regressors are weak.

5.5 Data

The data used in this chapter are from the same survey as used in chapter four and described in chapter three. The sub-sample employed in this analysis is composed of those individuals of single status who defined their main economic activity as either working for payment or profit in non-agricultural activities. Only those who classified themselves as full-time workers are included.

The sub-sample was allocated between the broad non-manual/manual occupational categories on the basis of the Census of Population Classification of Occupations (1981). Higher Professional, Lower Professional, Self-Employed and Managers, Salaried Employees, Intermediate Non-Manual Workers, and Other Non-Manual Workers (code nos. 2, 3, 4, 5, 6 and 7 respectively) were allocated to the non-manual sector. Skilled Manual Workers, Semi-Skilled Manual Workers and the Unskilled Workers (code nos. 8, 9 and X) were allocated to the manual collar sector.

The total number of observations for which no missing values were recorded was 2827. Of these, 1566 were non-manual and 1261 were manual workers. In terms of the males 568 were non-manual and 937 manual and for the females the comparable split was 998 and 324.

---

[43] A logit model could be used to obtain the predicted probabilities. However, in terms of this study the results are not substantially different under this alternative assumption.

The variables used in the estimation of both wage and occupational equations are as follows:

Wage :Net hourly wage expressed in logarithms.

Experience: Total labour force experience expressed in years.

Previous Experience: Experience prior to the current job expressed in years.

Education: Number of years in post-compulsory education. The school leaving age is fifteen.

Occupation: A (0,1) dummy variable assuming a value of 1 if the individual holds a non-manual job and 0 if a manual job is held. This variable serves as the dichotomous realisation of the latent dependent variable of equation (5.1).

Region of Schooling: A dummy variable adopting a value of 1 if the individual's region of schooling is in Dublin City or county and zero otherwise.

Firm Size: A set of three (0,1) dummies for the size of the firm the individual currently works in. The three dummies are for firms less than fifty workers, firms with greater than fifty but with less than four hundred workers and firms with greater than four hundred workers. In estimation the omitted dummy is firms with less than fifty workers.

Unemployment: This variable is calculated as the number of months an individual has spent unemployed since leaving full-time education.

Move Residence: A (0,1) dummy adopting a value of 1 if the individual changed residence to take their current job.

Father's Occupation: A set of two (0,1) dummies capturing the occupational status of the individual's father. The broad occupational categories are Non-Manual and Manual and the omitted category in estimation is agriculture.

A number of other variables were also used in estimation but failed to show a statistically significant effect. These include the number of jobs an individual held prior to the current one, the full set of region of schooling dummies and a set of industry dummies. The latter set of dummies in particular proved sensitive to slight alterations in the specifications and are thus excluded.

Finally, appendix A1 contains a set of summary statistics for the full set of variables used in the estimation.

5.6 Wage Equation Estimation

The sectoral wage equations could best be interpreted in terms of an integrated human capital/ compensating differentials explanation of wage determination. The standard human capital variables of schooling and post-schooling investments are present. The education variable is expressed in terms of years in post-compulsory education. As in chapter four the labour force experience variable is expressed in terms of two linear splines. The nodes used are based on less than or equal to four years labour force experience and strictly greater than four years labour force experience. Since three contrasting econometric methods are used in estimation statistically testing for the optimal nodes would prove complicated and for this reason is avoided. However, the four year split provides a more reasonable fit to the data for the sample used here than the five year split used in chapter four.

A dummy variable for whether the individual's residence of schooling was in Dublin city or county is also used in the wage equation and some justification for this particular variable must be provided. It could be argued that the residence of schooling variable implies more about the individual than the job the individual currently holds. Since both variables proxy different effects a strong argument exists for the inclusion of both. However, since the sample of workers are relatively young there exists a high correlation between residence of schooling in Dublin and current residence in Dublin and for the manual female workers the correlation is perfect[44]. Thus inclusion of both variables is vitiated for at least one of the wage equations. The approach adopted in this study is to include the residence of schooling dummy and bear the above caveats in mind when interpreting the coefficient estimates.

Two further variables that may be interpreted as loosely proxying job search variables are also included. One is a dummy variable for whether or not an individual changed residence to

[44] Regressions have been performed with both variables where possible. However, estimates proved imprecise due inevitably to the high correlations.

take up their current job and another captures the length of unemployment in months experienced by an individual since leaving full-time education. A set of firm size dummies designed to account for the effects of compensating differentials are included with the omitted reference group firms with less than fifty workers. A number of other variables that were also used in the analysis but to no effect were a set of industry dummies that proved very sensitive to alterations in the specifications and the number of previous jobs held by the individual. Estimates based on the use of these variables are thus not reported[45].

5.7 Wage Equation Estimates

Probit estimates for the reduced form occupational attachment equations are contained in appendix 5.A1 of this chapter. Comment on the occupational equations' results are reserved until discussion of the structural models.

Tables 5.1 and 5.2 contain OLS, IV and Heckman estimates for the male and female wage equations respectively. The coefficients on the experience variables appear relatively robust for both gender groups regardless of the econometric method used. The returns to labour force experience are, in general, greater in the early years in comport with human capital predictions and in the particular case of manual male workers greater returns to firm specific investments are observed.

The private rates of return to education appear more sensitive to the econometric method used and this is especially so for the female manual workers. Rates of 1.5% and 8.5% respectively for the Heckman and the IV techniques are recorded for the female manual category as compared to 3.6% for the OLS estimate. In both the IV and Heckman cases neither estimate is statistically significant at a satisfactory level which is in marked contrast to the manual male estimates which appear almost identical using IV or Heckman. The contrast may be explained to some extent by the small number of observations in the female manual category. The non-manual estimates for private educational returns for both sexes are much more in agreement and are resonant

[45] All of the empirical analysis was carried out on the LIMDEP econometric package with the exception of the calculations for the IV variance/covariance matrices which were calculated using the SAS package.

of similar findings in the human capital literature.

The estimated coefficients on the firm size dummies are, like the experience coefficients, relatively robust to the econometric technique. In most cases the returns are incremental with size. Though the manual female IV estimates are slightly out of line with the other estimates in terms of magnitude. The small number of observations in this category may be invoked to explain this particular phenomenon.

The coefficient on the schooling in Dublin dummy follows a similar pattern regardless of which estimator is used. As the discussion in section 5.6 indicated interpretation of this coefficient is difficult since it proxies not only individual attributes but also job attributes. However, the clear pattern that emerges for both sexes is the contrast in manual and non-manual wage effects that exists for individuals whose residence of schooling was in Dublin city or county. For manual workers from both sexes the effects are negligible in comparison to non-manual workers who record positive wage effects of well over 10% in most cases. This could be interpreted in terms of the large number of non-manual jobs in Dublin city or county and since the probit estimates suggest that young workers educated in Dublin are more likely to end up in non-manual jobs this need not emerge as a surprising result.

The unemployment variable included in the wage equations record some interesting results. At least two interpretations are possible for the results obtained for manual male workers. One interpretation of the positive wage effect is in terms of a premium for protracted job search. Another more plausible interpretation may be that employers do not use a manual male worker's duration of unemployment as a productivity proxy. The effects for non-manual males are statistically insignificant. Thus, for those young male workers in employment in the sub sample experience of unemployment does not possess a strong wage disadvantage. In contrast, most of the statistically significant female effects are negative in sign and operate through the non-manual sector. This a differential treatment of males and females in terms of unemployment and wages and may be interpreted as some indirect form of discrimination. Nothing of note is reported for the other job search variable, the change of residence dummy.

Table 5.1

Male Wage Coefficient Estimates.

Variable	OLS	std.error	IV	std.error	Heckman	std.error
Manual						
Constant	-0.2853***	0.0356	-0.3810***	0.0813	-0.3268***	0.0387
Exp.≤4 yrs.	0.1510***	0.0108	0.1750***	0.0195	0.1505***	0.0095
Exp.>4 yrs.	0.0568***	0.0081	0.0294	0.0203	0.0537***	0.0104
Education	0.0634***	0.0086	0.0438**	0.0209	0.0418***	0.0142
50 ≤ Firm < 400	0.1246***	0.0261	0.1171**	0.0568	0.1311***	0.0255
Firm ≥ 400	0.1421***	0.0258	0.2287***	0.0605	0.1330***	0.0266
Schooling in Dublin	0.0537**	0.0225	-0.0379	0.0666	-0.0021	0.0381
Unemployment(months)	0.0056***	0.0437	0.0079***	0.0020	0.0067***	0.0015
Move Residence	0.0391	0.0437	0.1407*	0.0729	0.0247	0.0510
Selectivity Bias	-	-	-	-	0.1621*	0.0854
Non-Manual	-	-	0.1407	0.2020	-	-
Constant	-0.1474***	0.0437	-	-	-0.1119	0.1265
Exp.≤4 yrs.	0.1003***	0.0120	0.0565*	0.0294	0.0998***	0.0122
Exp.>4 yrs.	0.0484***	0.0140	0.0870**	0.359	0.0480***	0.0141
Education	0.0683***	0.0104	0.0492***	0.0248	0.0646***	0.0150
50 ≤ Firm < 400	0.1388***	0.0337	0.1605	0.1081	0.1391***	0.0343
Firm ≥ 400	0.1987***	0.0309	0.0479	0.1102	0.1954***	0.0328
Schooling in Dublin	0.1242***	0.0267	0.1693**	0.0762	0.1158***	0.0391
Unemployment (months)	0.0032	0.0024	-0.0023	0.0064	0.0034	0.0034
Move Residence	0.0723*	0.0427	0.0099	0.0820	0.0697	0.0498
Selectivity Bias	-	-	-	-	0.0242	0.0808
Observations	1505		1505		1505	

***** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

Table 5.2

Female Wage Coefficient Estimates.

Variable	OLS	std error	IV	std. error	Heckman	std. error
Manual	-	-				
Constant	-0.0316	0.0633	-0.5117**	0.1987	-0.0981	0.1098
Exp.≤4 yrs.	0.0616***	0.0151	0.1067***	0.0251	0.0601***	0.0127
Exp.>4 yrs.	0.0151	0.0106	-0.0016	0.0259	0.0157	0.0118
Education	0.0362***	0.0126	0.0849*	0.0503	0.0157	0.0326
50 ≤ Firm < 400	0.1196***	0.0403	0.4865***	0.1695	0.1717**	0.0847
Firm ≥ 400	0.2820***	0.0418	0.6542***	0.1689	0.3072***	0.0550
Schooling in Dublin	0.0435*	0.0262	0.0270	0.0626	0.0229	0.0435
Unemployment(months)	0.0062***	0.0019	0.0054	0.0044	0.0019***	0.0007
Move Residence	0.1188**	0.0553	-0.0714	0.3396	0.1006	0.0863
Selectivity Bias	-	-	-	-	0.0666	0.0999
Non-Manual	-	-	-0.1209	0.0704*	-	-
Constant	-0.2298***	0.0361	-	-	-0.2357***	0.0534
Exp.≤4 yrs.	0.0883***	0.0092	0.0734***	0.0122	0.0885***	0.0084
Exp.>4 yrs.	0.0435***	0.0117	0.0588**	0.0277	0.0433***	0.0117
Education	0.0879***	0.0079	0.0686***	0.0140	0.0884***	0.0127
50 ≤ Firm < 400	0.1933***	0.0244	0.1036*	0.0571	0.1888***	0.0424
Firm ≥ 400	0.2714***	0.0223	0.2151***	0.0294	0.2698***	0.0255
Schooling in Dublin	0.1391***	0.0190	0.1575***	0.0281	0.1407***	0.0231
Unemployment (months)	-0.0055**	0.0023	-0.0018	0.0036	-0.0014**	0.0006
Move Residence	0.0115	0.0378	0.0456	0.0474	0.0124	0.0338
Selectivity Bias	-	-	-	-	-0.0103	0.0779
Observations	1322		1322		1322	

***** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

5.8 Endogeneity of Occupations

Two different approaches designed to test occupational exogeneity are examined. The first focuses on a modified statistical test based on the proxy variables of the Heckman procedure while the second uses the Hausman (1978) test as outlined in Duncan and Leigh (1985).

Melino (1982) provides a Lagrange multiplier test which is shown to be equivalent to the test of Heckman (1979) but possessing more desirable asymptotic properties. The LM test suggested is calculated as the square of the t-statistic associated with the proxy construct using the OLS variance-covariance matrix. This is consistent under the null hypothesis of occupational exogeneity in this case. The resultant test statistic is asymptotically distributed as χ^2 with one degree of freedom. In terms of the two-step Heckman procedure four independent χ^2 variates each possessing one degree of freedom are provided to statistically test the proposition of occupational exogeneity.

Hausman (1978) provides an alternative test for exogeneity based on the statistical comparison of the IV and OLS estimators. The test requires the comparison of an estimator that is consistent and efficient under the null hypothesis (of exogeneity) but inconsistent under the alternative (of endogeneity), *i.e.* the OLS estimator, with an estimator that is consistent under both hypotheses but inefficient, *i.e.* the IV estimator. The inefficiency of the IV estimator is due to the fact that the instrumental variables used may not be highly correlated with the corresponding independent variables. Consistency is thus bought at the cost of a high variance.

The Hausman test statistic is calculated as follows. Define

$$\hat{q} = \hat{\beta}_{IV} - \hat{\beta}_{OLS}$$

and

$$V(\hat{q}) = V(\hat{\beta}_{IV}) - V(\hat{\beta}_{OLS})$$

where V( ) the variance/covariance matrix of the estimator in question. Then the test statistic is given by

$$m = \phi' [V(\phi)]^{-1} \phi \qquad (5.15)$$

where m is distributed as χ^2 with k degrees of freedom where k is the number of parameters estimated.

Attention now turns to the test results. The results based on the LM tests of occupational exogeneity for the females in both the manual and non-manual sectors suggest little evidence of non-randomness in their allocation to either sector. In both cases the null hypothesis of occupational exogeneity is upheld by the data. Thus, the net hourly wage for a female with a given set of personal characteristics selected at random from the population into the manual sector is no different from what it would have been if selected into the non-manual sector.

The comparable LM test for non-manual males also records little evidence of selectivity bias. However, the LM test associated with the male manual sector fails to reject the null hypothesis (the test value is just outside the 5% level but comfortably inside the 10% level). The interpretation of this result needs to be explained not only in the context of statistical significance but also in terms of the direction that the selectivity bias operates. The signs on the estimated proxy variables provide information on the direction of the selectivity bias. As Lee (1978) points out the signs of the truncated effects can be analytically determined à *priori* by the second moments of the disturbances η_{mi}, η_{ni} and ϵ_i, (see Lee (1978) p.426). Due to the construction of the proxy variables (see (5.6) and (5.7) above) the positive coefficient implies negative selectivity or (truncation) in terms of manual jobs. The interpretation for this is that a male worker with a given set of characteristics selected at random from the population receives a lower wage if selected into the manual rather than the non-manual sector. In other words, on the basis of the male manual workers' coefficient estimates the wage distribution actually observed for male manual workers is over 16% lower than the mean wage offer distribution estimated for the population as a whole. Though the effect is only significant at a 10% level the result itself is large and requires some explanation. Chapter six using virtually the same sample suggests that the male manual selectivity bias effect operates through the skilled sector. This implies that a young male selected into the skilled sector earns a wage less than if selected into any other sector. This result is supported by the predictions of human capital theory which suggests that young workers sacrifice wages for training in return for greater life-cycle earnings. Though, in this study the manual category employed here is an amalgam of skilled, semi-skilled and unskilled workers the greater proportion of males (70%) are in the skilled category and the interpretation offered is regarded as reasonably plausible[46].

The Hausman test results are in slight contrast to those obtained using Melino's LM test. The χ^2 statistics are 7.82 and 8.88 for the male and female equations respectively implying that the null of exogeneity is upheld at any reasonable level of significance. Taking both sets of results together the data suggest little evidence of occupational endogeneity with the exception of the male manual category.

However, a number of caveats need to be made regarding the reliability of both sets of exogeneity results. Firstly, the LM test could be considered a test contingent on the assumption of normality being satisfied. If the assumption of normality is violated the LM test is invalid. Departures from normality will induce inappropriate inferences regarding the LM test. Chesher and Irish (1987) provide a set of easily computable diagnostics based on score tests for the probit and related models. Score tests for the null hypotheses of homoscedasticity, normality and of the information matrix identity are easily computable for the probit. In terms of the reduced form criterion function of (5.1) the null of normality is decisively rejected by the data[47]. The magnitude of this rejection brings the validity of the score test somewhat into question and as in the case of the linear regression model may to some extent be explained by the test's poor finite sample properties. Chesher and Spady (1988) examine this particular issue in the context of the linear regression model. If the score test results are valid and the distributional assumption of normality questionable Lee (1983) provides an alternative estimator that is consistent with more general distributional assumptions and suggests, for example, a logistic distribution[48]. However, this still possesses the limitation that a distributional assumption must be made. It is this distri-

[46] The χ^2 statistics for the males are 0.088 and 3.482 for the non-manual and manual sectors respectively and the comparable female estimates are 0.017 and 0.429 for non-manual and manual respectively. The associated critical values are 6.63, 3.84 and 2.71 for the 1%, 5% and 10% level of significance.

[47] χ^2 statistics of 656.6 and 846.5 are recorded for the male and female reduced form equations of (5.1) respectively. Both possess two degrees of freedom and the critical value at the 1% level is 9.21.

[48] Applying a logit to the reduced form of (5.1) makes little difference to the exogeneity results even if one accepts that the Lee test based on the J variable (like Heckman's) possesses less desirable asymptotic

butional problem that renders the IV estimator more appealing. However, the Hausman test possesses its own limitations. If the instruments are orthogonal to the regressor being instrumented the power of the test is zero and not rejecting the null when untrue is conceivable.

Therefore, both sets of results should be couched in terms of the above provisos. Since the LM tests could be taken as slightly in conflict with the Hausman tests any conclusion regarding exogeneity/endogeneity must remain relatively neutral.

5.9 Wage Differentials

The interpretation of the IV and Heckman based wage differentials using the "index number" approach is slightly different from OLS. The "index number" approach hinges on the fact that the regression "line" passes through the means of the data. Although, OLS satisfies this property the IV estimator doesn't. The IV based differentials are calculated in a similar manner to the OLS based differentials, *i.e.* by weighting mean characteristics by the appropriate IV coefficient estimates. The consequence of this, however, is that the explained and unexplained parts of the observed wage differential do not add up to the gross observed mean wage differential. Nevertheless, for purposes of comparison the IV based differentials are calculated in this manner.

In analysing wage discrimination it may appear more appealing to focus attention on the differential in wage offers that individuals drawn at random from the population face. A modification due to Reimers (1983) provides an alternative interpretation of the discrimination effect based on the differences in wages corrected for selectivity bias. The gross difference in observed wages for either occupational sector based on the two-step Heckman procedure is given by $\Delta\overline{W} = \Delta\overline{X}\hat{\beta}^m + \overline{X}^f\Delta\hat{\beta} + \hat{\theta}^m\overline{\lambda}^m - \hat{\theta}^f\overline{\lambda}^f$ (ignoring occupational subscripts). If that part of the observed wage differential attributable to selectivity bias is netted out the resultant differential, $\Delta\overline{W}^* = \Delta\overline{X}\hat{\beta}^m + \overline{X}^f\Delta\hat{\beta}$, represents the difference in wage offers. Thus, $\Delta\overline{W}^*$ is the observed wage differential adjusted for differences in selectivity bias between the sexes. $\overline{X}^m\hat{\beta}^m$ and $\overline{X}^f\hat{\beta}^f$ can be interpreted as providing consistent estimates of the wage offers for males and females respective-

properties than the Melino LM test.

ly, with mean characteristics, selected into either the manual or non-manual sectors. This wage offer differential, like OLS, may be broken down into explained and unexplained parts. From the point of view of discrimination, examining the gender wage offer differential provides an additional insight that focusing on the observed wage differential lacks. Thus, the Heckman based differential estimates reported are based on the observed wage corrected for selectivity bias, *i.e.* the wage offer.

The observed manual differential is quite small and suggests that on average male manual workers get 1.7% more than female manual workers. In contrast, the non-manual observed differential is considerably larger and suggests that non-manual males earn, on average, 9.5% more than their female counterparts. Tables 5.3 to 5.5 report OLS, IV and Heckman based explained and unexplained differentials in an attempt to establish how much of the gross differential is explained by characteristics and how much by differing coefficients[49].

The OLS estimates of table 5.3 reveal no evidence of a statistical difference between what males and females earn in the manual sector. To pre-judge tables 5.4 and 5.5 similar results are recorded for the IV and Heckman cases. However, as has already been alluded to, manual male workers may be sacrificing wages for training and any attempt to quantify wage based discrimination in a young sample of workers, with a high proportion of skilled manual workers, may be inappropriate. The problem with interpreting the manual differentials is also compounded by the fact that there is a small number of females in the skilled occupational category. Though there may be econometric advantages with the dichotomous occupational breakdown employed, the problem the categorisation creates for interpretation is clear.

The overall picture is slightly different when the unexplained non-manual differentials are examined. The OLS estimates of table 5.3 imply that males in the non-manual sector receive on average 5.9% more than females with comparable human capital and other characteristics. The differential in this case is statistically significant at the 1% level and could be interpreted as providing some evidence of wage discrimination. The comparable IV based estimate is considerably

[49] The standard errors reported are based on Stewart (1987) and their calculation is as in chapter four.

larger suggesting a male mark-up of over 16%. However, a cautious interpretation is suggested by its statistical insignificance. The gulf in estimates is surprising given the non-rejection of the null of exogeneity by the Hausman test. If the consistent and unbiased IV estimates are based on the correct treatment of occupations then the statistical insignificance could reflect the fact that the consistency is bought at the price of efficiency. The high variances recorded for the estimates may reflect the use of relatively poor instruments which would also bring the validity of the Hausman test results back into question. If because of the poor set of instruments used the Hausman test fails to detect endogeneity when present an explanation for the vast magnitude in estimates could be provided. Duncan and Leigh (1985) suggest that in the presence of endogeneity OLS based differentials should be different from IV based differentials since they represent the outcome of two conceptually different experiments. The former represents a differential conditional on an individual's occupational attachment and the latter represents a differential for an individual drawn at random from the population unconditional on occupational attachment.

Table 5.5 reports estimates of the explained and unexplained wage offer differentials using the Heckman procedure. The non-manual sector again indicates evidence of wage discrimination recording an estimate of 8.5% (significant at the 10% level). This could be interpreted as males in the non-manual sector receiving wage offers which are, on average, 8.5% more than what females are offered in the same sector with comparable endowments. Despite the slightly different interpretation the result is in line with the OLS based estimate which in the light of the LM test of the previous section is as one would expect. Since the selectivity effects for the female equations were both numerically small and statistically insignificant the comparability of the OLS and Heckman based differential estimates (though admittedly measuring two different concepts) is not too surprising.

Three salient conclusions emerge from this section. Firstly, no evidence of wage based discrimination exists in the manual sector. In the context of this sample of young workers this need represent no surprise. A large proportion of males in the manual sector are in the skilled category (70%) and in receipt of relatively low wages as firm specific human capital suggests. Thus, the failure to detect wage based discrimination may be itself concealing more insidious forms of discrimination in, for example, female access to skilled training. If one accepted the above conjecture, then, wider manual differentials can be anticipated with the passage of these young workers into the adult labour market.

Secondly, regardless of the econometric method used there is strong evidence of wage discrimination in the non-manual sector. The estimates range from between 6% to 16% and should be interpreted as more disturbing in view of the fact that over 75% of young females are in this particular sector and bearing in mind the single status and youth of the sample. In the latter context, in particular, the magnitude of the differential in the non-manual sector could be considered inordinately high. The long term implications that this suggests for the transition of these workers to the adult labour market should not be understated.

Thirdly, the low observed aggregate gender wage differential of 2.8% provides a misleading figure and disguises vaster gender differences in wages by occupational sectors however broadly defined. Invoking the manual/non-manual framework allowed a more worthwhile insight into the detection of unexplained gender wage differentials. This in itself may be interpreted as a vindication of the exercise undertaken.

Table 5.3

OLS Wage Differentials by Occupational Sector.

Sector	$\Delta\overline{W}$	$\Delta\overline{X}\beta^{m}$	$\overline{X}'\Delta\beta$
Non-Manual	0.0904	0.0327***	0.0576***
		(0.0036)	(0.0170)
Manual	0.0167	-0.0121	0.0288
		(0.0083)	(0.0189)

Asymptotic standard errors in parenthesis. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

Table 5.4

IV Wage Differentials by Occupational Sector.

Sector	$\Delta\overline{W}$	$\Delta\overline{X}\beta^{m}$	$\overline{X}'\Delta\beta$
Non-Manual	0.0904	0.0269**	0.1521
		(0.0106)	(0.0890)
Manual	0.0167	-0.0040	0.0044
		(0.0182)	(0.1002)

Asymptotic standard errors in parenthesis. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

Table 5.5

Heckman Wage Offer Differentials by Occupational Sector.

Sector	$\Delta\overline{W}^{o}$	$\Delta\overline{X}\hat{\beta}^{m}$	$\overline{X}'\Delta\hat{\beta}$
Non-Manual	0.1143	0.0323***	0.0819*
		(0.0022)	(0.0471)
Manual	-0.0076	-0.0175	0.0099
		(0.0112)	(0.1059)

$\Delta\overline{W}^{o}$ is the observed wage differential corrected for selectivity bias, *i.e.* the wage offer differential. Asymptotic standard errors in parenthesis. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

Table 5.6

Non-Manual/Manual Wage Differentials by Gender.

Gender	OLS	IV	Heckman
Male	0.0324	0.2242	0.1550
Female	0.0074	0.1027	0.0994

5.10 Structural Model Estimates

Though the estimation of the structural model does not represent the major thrust of this chapter estimates are reported for completeness sake. The estimates themselves may be viewed of some assistance in shedding light on the major determinants of occupational attachment for young workers and in particular the role played by the non-manual/manual wage gap. Variables used in the estimation of the structural occupational model comprise all those used in estimating the reduced form of (5.1) with the exception of identifying restrictions provided by the firm size dummies and the inclusion of an appropriate experience variable and the wage gap term. In contrast to the reduced form the experience variable used in the structural model is experience prior to the current job. It should be recalled that full experience entered the reduced form through substitution of the wage equations into the structural occupational equation.

The wage gap variable is estimated as the difference between what an individual would earn if non-manual and what the same individual would earn if manual. Since three estimators are used to estimate the wage equation parameters, three separate wage gap variables suggest themselves implying the estimation of three separate structural occupational models for each gender group.

The structural occupational model is estimated by a maximum likelihood probit estimator. In order to facilitate interpretation the maximum likelihood estimates are transformed into their marginal effects. Thus, the effect of the k^{th} variable on the probability of non-manual attachment is given by:

$$\frac{\delta P}{\delta \Gamma_k}(Y_i = 1) = \mu_k \phi(\Gamma' \mu) \tag{5.16}$$

where Γ is the matrix of explanatory variables in the structural model, $\phi(\,)$ is the standard normal density function and μ_k is the k^{th} coefficient of interest. The results reported in tables 5.7 and 5.8 represent the marginal effects with associated maximum likelihood t-statistics in brackets.

As table 5.6 indicates the means of the occupational wage differentials vary surprisingly depending on which estimator is employed. The mean of the OLS estimate for the non-manual/manual differential is 3.3% for the males and just under 1% for the females. In controlling for endogeneity the comparable figures using the consistent IV estimates are 25.1% and 10.8% and using the consistent Heckman estimates 16.8% and 10.4%. A key question, therefore, is the extent to which the structural model estimates are sensitive to the wage gap term used in estimation. Inspection of the results reported in tables 5.7 and 5.8 indicates that most of the explanatory variables determining non-manual attachment appear relatively robust to the wage gap variable used. Where the contrast in estimates actually occurs is, in fact, in terms of the coefficients on the wage gap variables. However, before discussing the wage gap estimates themselves a brief discussion initially focuses on the effects of the other explanatory variables on non-manual male occupational attachment.

Possession of previous experience is negligible in terms of influencing non-manual attachment. For young workers this need not be surprising since employers selection of non-manual young workers is more likely to be on the basis of their educational qualifications than on previous experience. In support of this it is clear that the more post-compulsory education one has the stronger is the probability of non-manual attachment. On average, the estimates suggest that one additional year of education increases an individual's probability of non-manual attachment by over 8%. Having received one's secondary education in Dublin city or county also has associated with it a remarkably strong effect. However, this effect may be capturing the strong correlation between Dublin and non-manual jobs. In view of the youth of the sample and the consequent high correlation between residence of schooling and current residence it seems more likely that this strong relationship expresses more about non-manual jobs being Dublin based than it does about the individuals themselves. The well determined negative coefficient on the duration of unemployment is consistent with the belief that white collar employers use unemployment as either a productivity index or a screening device in their allocation of workers to non-manual jobs. Finally, the parental background variables assumed proxied by father's occupation appear to suggest that, for males at least, possessing a non-manual father relative to an agricultural father increases the probability of such an individual's non-manual attachment by over 20%. Comment on the wage gap term estimates is reserved until later.

All the above comments concerning the male estimates could be repeated for the female estimates of table 5.8. Prior experience doesn't show statistically and the educational effect appears even more pronounced for the female equation. The Dublin schooling coefficient and the unemployment variables both record effects similar to those reported for the male equation. However, the one interesting feature of the female equation, apart from the wage gap variable, is the parental background effects. They are, in general, either statistically insignificant or negative. This suggests that the female progeny of the agricultural sector *i.e.* daughters of farmers fare at least as well as, if not better than, other groups in securing non-manual employment. This effect could be explained by a combination of social conditioning and/or lack of opportunity within the agricultural environment for females.

Attention now turns to an examination of the coefficients on the wage gap terms of the structural models. The male terms are all statistically insignificant and with the exception of the IV based wage gap variable large standard errors are recorded. Thus, the tentative conclusion offered for males is that wage gaps are not as important in determining non-manual attachment as, for example, is education, residence of schooling, unemployment or parental background. The female estimates present a degree of ambiguity. For the three wage gap variables in the female equation, all are statistically significant at the 5% level or better. However, while the OLS and the Heckman based estimates record negative effects, the IV based wage gap variable records a strong positive effect. While all those variables that are important for the males are also as important for the females the conflict in signs in terms of the wage gap variables and their statistical significance is worrying. No convincing interpretation for this particular result can be offered.

Table 5.7

Marginal Effects for the Structural Male Equations

Variable	OLS based	IV based	Heckman based
Constant	-0.3930***	-0.3695***	-0.3950***
	(12.337)	(10.357)	(11.332)
Previous Exp.	0.0154*	0.0137*	0.0159**
	(1.936)	(1.856)	(2.016)
Education	0.0815***	0.0822***	0.0806
	(9.446)	(10.113)	(7.889)
Schooling in Dublin	0.1614***	0.1776***	0.1576***
	(5.477)	(6.254)	(4.542)
Move Residence	0.0855*	0.0618	0.0842*
	(1.712)	(2.664)	(1.674)
Unemployment (in months)	-0.0047	-0.0057***	-0.0047***
	(-2.322)	(2.664)	(2.261)
Father Non-Manual	0.1913***	0.1910***	0.1911***
	(5.755)	(5.754)	(5.749)
Father Manual	0.0522*	0.0519	0.0522*
	(1.652)	(1.643)	(1.652)
Wage Gap (logs)	-0.0060	-0.1038	0.0259
	(0.0330)	(1.413)	(0.140)
Dep.Var. (mean)	0.377	0.377	0.377
χ^2	269.13***	271.12***	269.15***
ρ^2	0.135	0.136	0.135
Observations	1505	1505	1505

Values in parentheses are |t| values. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. χ^2 tests for the overall significance of the slope coefficients and ρ^2 is the McFadden pseudo-R^2.**

Table 5.8

Marginal Effects for the Structural Female Equations

Variable	OLS based	IV based	Heckman based
Constant	-0.0827**	-0.0808***	-0.0112
	(2.572)	(2.642)	(0.365)
Previous Exp.	0.0220***	-0.0041	0.0100
	(2.904)	(0.585)	(1.252)
Education	0.1477***	0.1205***	0.1297***
	(13.007)	(13.432)	(8.450)
Schooling in Dublin	0.2135***	0.0736***	0.1595***
	(7.316)	(3.072)	(4.755)
Move Residence	-0.0369	0.0658	0.0517
	(0.690)	(1.396)	(0.961)
Unemployment (in months)	-0.0220***	-0.0047**	-0.0035***
	(7.476)	(2.344)	(4.051)
Father Non-Manual	0.0008	0.0227	0.0051
	(0.026)	(0.746)	(0.162)
Father Manual	-0.0836	-0.0626**	-0.0838***
	(3.101)	(2.266)	(2.936)
Wage Gap (logs)	-1.1004	0.4642***	-0.4142**
	(6.599)	(9.040)	(2.093)
Dep.Var. (mean)	0.755	0.755	0.755
χ^2	311.76***	354.90***	279.91***
ρ^2	0.212	0.240	0.185
Observations	1322	1322	1322

Values in parentheses are |t| values. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. χ^2 tests for the overall significance of the slope [illegible] and ρ^2 is the [illegible]**

5.11 Conclusions

A key issue addressed in this chapter has been the effect of occupational endogeneity on both gender and occupational wage differentials. Two contrasting econometric methods were employed to control for the potential endogeneity and both produced contrasting results. Use of the IV procedure with the associated Hausman test provided no statistical support for the endogeneity proposition in either the male or the female equations. In contrast the Heckman procedure provided some evidence of sample selectivity in terms of the allocation of young male workers to the manual sector. The caution expressed in interpreting these results was prompted by possible departures from assumed normality in the Heckman procedure and by the orthogonality of instruments and regressors in the IV case. It has to be accepted that the rationalisations presented were conjecture and not based on convincing evidence one way or the other. Nevertheless, all the estimators agreed that there is little statistical evidence of an unexplained gender wage differential in the manual sector and all agreed that the converse was the case for the non-manual sector. However, disagreement was recorded in terms the magnitude of the latter non-manual effect with estimates ranging between just over 5% to under 16%. The finding for the non-manual sector is made all the more interesting by the fact that 75% of young females work in this particular category. Thus, it could be argued that in the non-manual sector females languish in lower paid jobs in comparison to their equally qualified male counterparts. It could further be tentatively suggested that the female wage disadvantage lies in the fact that they do not secure the more senior jobs in this sector. This could be due to the fact that career ladders are shorter and promotional prospects lower for females. The fact that such a differential should exist in a sample of workers who are single and relatively young is disturbing not least for the long term implications it suggests.

It is, in the light of the above discussion, difficult to rank in order of preference the most appropriate estimation procedure, and hence, the most convincing differential estimate. If one is prepared to accept that occupational endogeneity is not an issue in this study and invoke the weak statistical test results as supportive of this view, then, the OLS estimates may be interpreted as relatively sound. However, decisive rejections of the normality assumption in the reduced form occupational attachment equations bring into question the Melino LM test. The possible weak correlation between the instruments and the corresponding independent variables may cast doubt on the reliability of the Hausman test. Given both these caveats no decisive ranking of estimation procedures is possible or prudent.

Structural models of occupational attachment were also estimated and some robust findings were detected. Education, residence of schooling, unemployment and parental background all played a role in one way or another in the determination of male and female occupational attachment. However, estimates of the male wage gap effect proved to be statistically insignificant with the comparable effect for females proving sensitive to the manner the wage gap variable was calculated. On balance it could be concluded that the effects of the non-manual/manual wage gaps are of less importance to young workers in the determination of occupational attachment than, for example, is education, parental background and unemployment.

Finally, it is clear from the preceeding analysis that the occupational segregation of females may explain some part of their wage disadvantage. It's clearly difficult to establish the gender wage effects that originate through occupational segregation on the basis of framework used in this chapter. Greater insight into the segregation effect is likely to come from the use of a finer occupational classification. Chapter six outlines a methodology designed to achieve this objective and quantify the gender wage effects associated with occupational segregation.

Appendix 5.A1

Table 5.A1

Reduced Form Male and Female Probit Estimates

Variable	Male	Female
Constant	-1.2403***	0.3767**
	(9.334)	(2.260)
Exp.≤4 yrs.	0.0217	0.0396
	(0.657)	(0.969)
Exp.>4 yrs.	0.0366	-0.0205
	(1.003)	(-0.467)
Education	0.2487***	0.4514***
	(9.600)	(11.780)
50 ≤ Firm < 400	-0.1135	-1.246***
	(1.254)	(11.879)
Firm ≥ 400	0.1066	-0.5452***
	(1.231)	(4.585)
Schooling in Dublin City & Co.	0.4744***	0.5166***
	(5.911)	(5.028)
Unemployment(months)	-0.0130**	-0.0373***
	(2.109)	(4.381)
Move Residence	0.2252	0.2915
	(1.456)	(1.406)
Father Non-Manual	0.5895***	0.1170
	(5.755)	(0.884)
Father Manual	0.1547	-0.2765**
	(1.589)	(2.307)
Observations	1505	1322

Values in parentheses are |t| values. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.

Chapter Six

Gender Wage Discrimination and Occupational Segregation

6.1 Introduction

The inclusion of exogenous occupation controls ignores, as Brown *et al.* (1980) point out, the existence of any potential discriminatory factors that may impinge on an individual's access to certain occupations and treats the given occupational distribution as justified. The authors argue for a more integrated approach to calculating the wage effects of gender discrimination and suggest a modification to the "index number" approach that allows for a decomposition of the gender wage difference into explained and unexplained occupational components. The approach merges aspects of the wage differentials literature (examples of which are cited in chapter two) and the occupational attachment literature (an example of which is provided by Schmidt and Strauss (1976)). The methodology allows not only for a decomposition of the gender wage differential into explained and unexplained parts within occupations (the intra-occupational effect) but also allows for gender wage differential effects that operate through differing occupational attainment (the inter-occupational effect). This latter effect can be further decomposed into effects due to differing characteristics and to those due to differing coefficients. This last term provides an estimate for the wage effects of occupational segregation. Brown *et al.* (1980) and Miller (1987) have both employed this methodology in empirical work for the US and UK respectively since it allows not only for a direct treatment of occupational differences but also for a more sensitive treatment of occupational wage differentials. As a consequence it is argued a more fruitful insight into discrimination effects is provided.

The methodology requires two distinct and separate steps. Firstly, the estimation of an occupational attachment equation in order to predict and simulate male and female occupational distributions and secondly the estimation of separate occupational wage equations. It is the latter step that provides for the isolation of a "cleaner" occupational wage discrimination effect. However, as chapter five pointed out the problem posed by the estimation of separate occupational wage equations relates to the possible existence of some selection process that determines the observed occupational sample. If the disturbance terms in the occupational wage equations are correlated with the disturbance term in the occupational selection equation then conventional estimation techniques provide biased and inconsistent parameter estimates. This has clear implications for the estimated discrimination effects. Methods designed to correct for such selectivity bias in a dichotomous context have been suggested in the literature by Heckman (1976 and 1979) and were applied in an occupational context in chapter five. However, little attention has focused on the problem as posed in the more valid polychotomous occupational framework[51].

The clear deficiency of the Brown *et al.* (1980) and the Miller (1987) papers lies in a failure to take account of the potential selectivity bias in occupational wage equations. An objective of this chapter is to integrate more effectively the information derived in the estimation of an occupational attachment equation with the occupational wage equations in order to correct for the potential presence of selectivity bias in the latter. An approach first outlined by Lee (1983) and used by Trost and Lee (1983) and Dolton, Makepeace and Van der Klaaw (1987) is employed here.[52] Though the "index number" decomposition modified by Brown *et al.* (1980) and used by Miller (1987) is contingent on the use of OLS an adjustment to take account of selectivity bias due to Reimers (1983) used in chapter five is also employed in this chapter.

The following section demonstrates how the observed mean wage differential may be decomposed into intra- and inter-occupational components. Section 6.3 explains the modelling and econometric methodology and section 6.4 provides a brief description of the data set to be used. Sections 6.5 and 6.6 report and comment on the occupational and wage equation estimates Sections 6.7 and 6.8 calculate the empirical components of the wage differential as outlined in

---

[51] Hay (1980) and Dolton, Makepeace and Van der Klaaw (1987) provide an exception in this regard.

[52] Though Dolton, Makepeace and Van der Klaaw (1987) examine the issue of selectivity bias in male and female wage equations an explicit examination of the effects of selectivity bias on the discrimination coefficient is avoided.

section 6.2. Section 6.9 offers some conclusions.

6.2 Decomposing the Wage Differential

Following Brown *et al.* (1980) the mean gender wage differential can be decomposed into explained and unexplained inter- and intra-occupational components. The male and female occupational wage equations may be expressed as

$$W_j^m = Z_j^m \beta_j^m + \varepsilon_j^m \tag{6.1}$$

$$W_j^f = Z_j^f \beta_j^f + \varepsilon_j^f \tag{6.2}$$

where the superscripts m and f denote respectively male and female, W_j^k (where k = m,f) denotes the wage for the j^{th} occupation expressed in logs. Z_j^k denotes a vector of the standard productivity variables for the j^{th} occupational category, β_j^k a vector of unknown parameters for the males and females and ε_j^k is an error term assumed to satisfy the standard set of assumptions.

The "index number" approach may be used to decompose the occupational wage differential into portions attributable to differing coefficients given the same endowments and to differing endowments given the same wage structure. To obtain the overall mean wage differential in this context the above components must be weighted by the sample proportions of males and females in each occupation. Denoting these sample proportions by P_j^m and P_j^f for the males and the females respectively, the overall mean wage differential may be written as[53]

$$\Delta\overline{W} = \overline{W}^m - \overline{W}^f = \sum_{j=1}^{M} (\overline{P}_j^m \overline{W}_j^m - \overline{P}_j^f \overline{W}_j^f) \tag{6.3}$$

However, information regarding the effect of occupational differences can also be incorporated into the above decomposition. Manipulating some of the terms in the above equation allows (6.3) to be re-written as

$$\sum_{j=1}^{M} \overline{P}_j^f (\overline{Z}_j^m \hat{\beta}_j^m - \overline{Z}_j^f \hat{\beta}_j^f) + \sum_{j=1}^{M} \overline{Z}_j^m \hat{\beta}_j^m (\overline{P}_j^m - \overline{P}_j^f) \tag{6.4}$$

[53] Here bars denote means, carats estimates and M equals the number of occupational categories.

Thus, the last term in expression (6.4) controls for the occupational distribution of women and captures the effect of occupational differences on the wage differential. Wages can be decomposed in the normal way assuming that, in the absence of discrimination, the male wage and occupational structure prevails. Expanding (6.4) further allows for the decomposition of the grand mean wage differential into its four component parts

$$\Delta\overline{W} = \overline{W}^m - \overline{W}^f = \sum_{j=1}^{K} \overline{P}_j^f \overline{Z}_j^f \Delta\beta + \sum_{j=1}^{K} \overline{P}_j^f \beta_j^m (\overline{Z}_j^m - \overline{Z}_j^f) + \sum_{j=1}^{K} \overline{W}_j^m (\overline{P}_j^m - \hat{\overline{P}}_j^f) + \sum_{j=1}^{K} \overline{W}_j^m (\hat{\overline{P}}_j^f - \overline{P}_j^f) \tag{6.5}$$

where $\Delta\beta = \beta_j^m - \beta_j^f$ and $\hat{\overline{P}}_j^f$ is the proportion of females in the sample that would be in the j^{th} occupation if females were confronted by the same distribution of occupational opportunities as males. Thus, the grand mean wage differential may be decomposed into four constituent parts. The first and second terms are the unexplained and explained gender differentials in wages. The third term represents that part of the overall mean wage differential attributable to the explained allocation of workers to the given occupational categories. The fourth and final term represents that part of the differential in the mean wage due to unexplained gender differences in the structure of occupational attainment. This may be interpreted as the effect of occupational segregation on the gender wage differential.

It should be pointed out that the results obtained are contingent on what assumption is made regarding the wage and occupational structure in the absence of discrimination[54]. For the purposes of this analysis the structure prevailing in the absence of discrimination is assumed to be fully described by the male structure. As already mentioned in chapter four this assumption does not appear too offensive. Brown *et al.* (1980) illustrate that the "index number" problem also applies in terms of the assumed occupational structure with the results again hinging on which structure is assumed to prevail in the absence of discrimination. A taste based model of discrimination can also lead to occupational segregation of females holding identical characteristics as males. If, again, one assumes males are not subject to the exercise of an employer's[55]

[54] This is the well known "index number" problem a discussion of which appears in chapter four.

discriminatory power then the male occupational distribution can be assumed to reasonably approximate a non-discriminatory occupational distribution.

The $\hat{\beta}$ coefficient estimates for each occupational category presented in this chapter are obtained using the conventional OLS estimator and using a consistent estimator due to Lee (1983). The selectivity bias correction, as Reimers (1983) shows, has implications for the observed gender wage gap at both occupational and aggregate levels. Though all four components of (6.5) can be isolated using the consistent estimator their interpretation is slightly different from OLS as was explained in chapter five.

6.3 Econometric Methodology

The theoretical background to occupational attachment is briefly sketched and follows in spirit the methodology outlined in both Hill (1983) and Trost and Lee (1983). Occupational attachment may be viewed in a utility based framework. Each individual is assumed to select from M mutually exclusive categories. The individual is further assumed to compute the utilities attainable from each category and choose that one which provides the maximum utility level.

More conveniently it is possible to express the maximum attainable utility for each of the M alternatives in terms of indirect utility functions. For the j^{th} occupational category, for instance, this may be expressed as

$$\mathbf{V}_j = \mathbf{V}(\mathbf{w}_j, \mathbf{Y}, \mathbf{T}, \mathbf{K}_j, \mathbf{R}) \qquad (6.6)$$

where,

$\mathbf{w}_j$ is the wage offer associated with occupation j

$\mathbf{Y}$ is non-labour income,

$\mathbf{T}$ is the endowment of time,

$\mathbf{K}_j$ is a vector of job characteristics associated with occupation j,

and $\mathbf{R}$ is a vector of exogenous variables.

[55] The analysis is equally valid in terms of an employee's or in some cases a consumer's taste for discrimination.

The utility based framework need not be interpreted as being inconsistent with labour market discrimination. For example, w_j and/or the vector K_j may differ across gender due to, for example, an employer's taste for discrimination. Lower wage offers and/or unfavourable job characteristics may reduce a female's indirect utility and hence her willingness to select given occupations.

The indirect utility function may be decomposed into stochastic and non-stochastic parts. If V_{ji} is the maximum utility attainable for individual i if occupation j is chosen then the indirect utility function may be expressed as

$$V_{ji} = S_{ji} + u_{ji} \tag{6.7}$$

The probability that the i^{th} individual chooses the j^{th} occupational category is given by

$$P_{ji} = \Pr\left[V_{ji} > V_{ki},\ \text{for } k \neq j,\ j = 1, \ldots\ldots M \right] \tag{6.8}$$

or alternatively,

$$P_{ji} = \Pr\left[S_{ji} - S_{ki} > u_{ki} - u_{ji},\ \text{for } k \neq j,\ j = 1, \ldots\ldots M \right] \tag{6.9}$$

Assuming the stochastic components have independent and identical Weibull distributions then the difference between the error terms $(u_{ki} - u_{ji})$ has a logistic distribution and the resultant model is the multinomial logit model due to McFadden (1973). As is obvious from the above only binary comparisons are involved and this follows from the strong behavioural assumption of the independence of irrelevant alternatives which gives the logit model its form.

For estimation purposes if S_{ji} is replaced by $X_i \gamma_j$ then the multinomial logit model may be expressed as

$$P_{ji} = \frac{\exp(X_i \gamma_j)}{\sum_{j=1}^{M} \exp(X_i \gamma_j)} \tag{6.10}$$

where X_i is assumed to capture all the relevant demand and supply effects contained in the indirect utility function and γ is vector of unknown occupational coefficients. Schmidt and Strauss (1976) and Brown *et al.* (1980) employed this particular model in estimating occupational attachment equations. Occupation is treated as a categorical, unordered, discrete polychotomous variable and the logistic approach is used to estimate the impact of a vector of explanatory variables on the probability of being in a particular occupation relative to another. The estimation of a multinomial logit model of occupational attainment allows prediction of an individual's occupational level on the basis of that individual's set of personal characteristics.

Miller and Volker (1985) suggest advantages for the use of an ordered probit approach and Miller (1987) uses this approach in an occupational application. However, for the purposes of this chapter the use of such an ordered approach is avoided for two reasons. Firstly, the sequential ranking of occupations should be on the basis of life-cycle earnings. In terms of the young workers used in this sample estimation of life-cycle earnings is not possible and to the author's knowledge no additional evidence on this particular subject is available for Ireland. Secondly, use of the ordered probit approach may possess greater advantages if the focus of attention (as in the Miller and Volker (1985) case) is vertical occupational mobility. In the context of this chapter the occupational mobility of interest is of the horizontal kind and this allows the unordered framework provided by the multinomial logit to be exploited.

A reduced form equation is assumed which describes the interaction of the relevant demand and supply conditions in the labour market and determines an individual's attachment to a certain occupation. Because of the reduced form nature of the estimating equations it is not possible to provide unambiguous interpretations for the coefficient estimates in terms of explicit demand or supply side effects. An eclectic theoretical view should be adopted in the interpretation of the coefficient estimates. In terms of (6.10) above only the parameters of $M-1$ of the M occupational categories can be identified. The following normalisation $\sum \gamma_m = 0$ is used in estimation and (6.10) becomes

$$\mathbf{P}_{ij} = \frac{\exp(\mathbf{X}_i \gamma_j)}{1 + \sum_{j=1}^{M-1} \exp(\mathbf{X}_i \gamma_j)} \qquad (6.11)$$

Alternatively the above expression may be expressed in terms of the log odds of being in a certain occupational category and this is a function that is linear in its parameters and is given by (6.12).

$$\log\left[\frac{P_{ij}}{P_{iM}}\right] = X_i \gamma_j \tag{6.12}$$

A dummy variable is used to define the event of an individual being in a certain occupation. $v_{ij} = 1$ if the i^{th} individual falls into the j^{th} category and $v_{ij} = 0$, otherwise. In this case the log likelihood function is given by

$$\log L = \sum_{i=1}^{N} \sum_{j=1}^{M} v_{ij} \log P_{ij} \tag{6.13}$$

where N equals the number of observations in the sample. Maximum likelihood methods are used to estimate (6.13). As pointed out above in estimation the parameters of the M^{th} occupational category are normalised to zero. The interpretation of the estimated multinomial coefficients are therefore in relation to this omitted category. Furthermore, the inclusion of intercept terms in the multinomial logit model ensures that the mean of the predicted probabilities equals the means of the actual probabilities. This is important in terms of the "index number" decomposition.

The next step is to use the information concerning occupational attachment in the estimation of the occupational wage equation. If one starts by assuming that the market wage in the j^{th} occupation is given by

$$W_j = Z_j \beta_j + e_j \tag{6.14}$$

where

W_j is the logged market wage for the j^{th} occupation,

Z_j is a vector of exogenous variables assumed to determine the wage in the j^{th} occupation,

β_j is a vector of unknown parameters,

and e_j is an error term for which the usual properties are assumed satisfied.

If a systematic process governs the observation of the j^{th} sample of wages and if the error term in that process is correlated with e_j then the application of OLS to the above equation leads to biased and inconsistent coefficient estimates. Following Lee (1983) the wage equation to be estimated may be modified to take into consideration the effects of this occupational selectivity bias. As Lee (1983) shows the wage equation conditional on category j being chosen is

$$W_j = Z_j\beta_j - \sigma_j\rho_j\frac{\phi(J(X\gamma_j))}{F(X\gamma_j)} + \zeta_j \tag{6.15}$$

where

ϕ is the standard normal density function,

J is a strictly increasing transformation that transforms the random variable associated with the occupational attachment equation into a standard normal variate where $J = \Phi^{-1}F$ where Φ is the standard normal distribution function and F is the probability distribution function. σ_j is the standard error of the disturbance term ε_j, and ρ_j are the correlations between ε_j and the error term from the occupational attachment equation for the i^{th} individual.

Estimation is carried out in a two step framework analogous to the Heckman procedure employed in chapter five. Firstly, maximum likelihood estimation is used to obtain estimates for $\hat{\gamma}_j$ from (6.11). Then, these estimates are inserted into the wage equation of (6.15) which may be re-written as

$$W_j = Z_j\beta_j - \sigma_j\rho_j\frac{\phi(J(X\hat{\gamma}_j))}{F(X\gamma_j)} + \zeta_j \tag{6.16}$$

where $F(X,\hat{\gamma}_j)$ are the predicted probabilities from the multinomial logit model of (6.11). More conveniently this equation may be expressed as

$$W_j = Z_j\beta_j + \theta_j\lambda_j + \zeta_j \tag{6.17}$$

where everything is as above with the exception of

$\theta_j = \sigma_j\rho_j$,

and $\lambda_j = -\frac{\phi(J(X\hat{\gamma}_j))}{F(X\gamma_j)}$. Consistent estimates for the j sector's wage are then obtained by the application of OLS to the above equation (6.17). The disturbance terms ζ_j are obviously heteroscedastic and Lee *et al.* (1980) outline an appropriate variance/covariance matrix in this regard[36]. However, this proved computationally difficult to calculate in the context of this chapter and so the White (1980) heteroscedastic consistent variance/covariance matrix is reported below for the occupational wage equations. Though the White variance/covariance matrix corrects for heteroscedasticity in the regression model it doesn't take into consideration the fact that a predicted

[36] See Chapter Five, section 5.3.

selectivity bias term is used in estimation. Nevertheless, though the variance/covariance matrix reported may be inappropriate differences are not anticipated to be large and as Maddala (1983) points out little is even lost in the use the OLS variance/covariance matrix. However, for the purposes of this chapter the White (1980) consistent variance/covariance matrix estimates are reported for both the OLS and the selectivity bias corrected occupational wage equations of (6.14) and (6.17) respectively.

The inclusion of the selectivity bias term has clear implications for the "index number decomposition". The modification suggested by Reimers (1983) and used in chapter five will also be used in this chapter. Thus, in summary, occupational attachment equations and wage equations are estimated for each gender. On the basis of the male occupational equation estimates female occupational distributions are simulated in order to obtain a handle on the occupational segregation effects. In addition female wages will be simulated on the basis of male wage structures for each occupational category to ascertain explained and unexplained wage effects within occupations. The analysis is presented for both observed wages and the wage offers associated with the consistent estimator.

In terms of the dependent polychotomous variable five relatively broad occupational categories are assumed. These are

(a) Professional and Managers,

(b) Clerical and Intermediate Non-Manual,

(c) Other Non-Manual,

(d) Skilled,

(e) Semi and Unskilled.

The five-way categorisation is dictated by the need to have a sufficient number of observations in all the relevant groups. A finer classification would lead to a reduction in the number of observations in particular estimating cells and would place the results in a somewhat dubious light. Too few females in the semi-skilled occupational category prevented a wider categorisation. However, it is felt that the classification used is broad enough to allow for some confidence in the estimation results and fine enough to examine the issue of within occupation wage discrimination.

Finally, the omitted occupational category in terms of estimation is the Semi and Unskilled category. Thus all the occupational equation coefficients should be interpreted in relation to this particular category.

6.4 Data

The number of individuals used in the analysis in this chapter is the same as chapter five, however, some of the variables used are slightly different. The sub-sample was allocated across the five occupational categories outlined in section 6.3 on the basis of the Census of Population Classification of Occupations (1981). Higher Professional, Lower Professional, Self-Employed, Managers, Salaried Employees were allocated to (a), Intermediate Non-Manual Workers to (b), and Other Non-Manual Workers were allocated to (c). Skilled Manual Workers were allocated to (d) and Semi-Skilled Manual Workers and the Unskilled Workers were allocated to (e).

The full set of variables used in the estimation of the reduced form occupational equations for both gender groups are

(i) An education variable defined in terms of the number of years spent in post-compulsory education.

(ii) A previous experience variable defined as the time spent working in jobs prior to the current one. The unit of measurement is years.

(iii) A set of residence of schooling dummies controlling for the area of an individual's last school prior to leaving compulsory education. The three estimated dummies are Dublin city and county, the east and midlands and the southern region with the omitted category schooling in the north-west.

(iv) A duration of unemployment variable defined in terms of the aggregate number of months of unemployment experienced by an individual since leaving school.

(v) A change of residence dummy set equal to 1 if the individual changed residence to take up their current job.

(vi) A set of Father's occupational dummies designed to capture parental background influences on occupational uptake. Two such dummies are defined, one for non-manual and another for the manual category. The agricultural category is treated as the reference category in estimation.

The occupational wage equations are estimated using variables (i), (iv) and (v) above in addition to

(vii) A full experience variable defined in terms of two linear splines with a four year split.

(viii) A set of two firm size dummies (see chapter five).

(ix) A set of four current region of residence dummies. The four estimated dummies are Dublin city and county, the east, midlands and the southern region with the omitted category current residence in the north/west[57].

A number of other variables were also used in estimation but to little statistical effect. Industry dummies were used in the wage equations but proved sensitive to alterations in the specification and thus are not included. For the occupational equation Father's occupation was broken down into a finer classification but some of the estimated coefficients possessed high standard errors. The number of jobs held by the individual since leaving school was also used in both occupational and wage equations but again to little effect.

The total number of observations for which no missing values are recorded is 2827, of which 1505 are male and 1332 female. Appendix A1 contains a set of summary statistics for the full set of variables used in the estimation.

6.5 Occupational Equation Estimates

The maximum likelihood multinomial logit estimates for males and females appear in tables 6.1 and 6.2 respectively with their associated asymptotic standard errors. The coefficient estimates themselves are not amenable to a ready interpretation. Schmidt and Strauss (1976) point out that the coefficients on the explanatory variables for the relevant categories present the

[57] For most of these variables a more in depth description is available in chapter five.

difference in coefficients between the category in question and the omitted category. It may be convenient to rank the estimated coefficients in order of size. Thus, the larger the coefficient estimate the greater is the impact of the associated explanatory variable on being in a given category.

The male coefficient estimates are broadly in line with one's priors. The education coefficient increases in size moving up the occupational ladder and reserves its largest effect, not surprisingly, for the professional category. The family background variables, in general, record statistically significant effects. Possessing a non-manual father relative to the agricultural reference group enhances one's probability of attachment to all occupational categories. The effect increases as one moves up the occupational categories providing the strongest effect for the professional category. A strong significant relationship between Dublin schooling and non-manual occupational destination for young male workers is also noted. This may be related to the fact that schooling in Dublin is highly correlated with current residence in Dublin which itself is highly correlated with the availability of non-manual jobs. The duration of an individuals unemployment since leaving school has a marked negative effect on attachment to all of the occupational categories. The largest effects recorded are for the higher non-manual categories. Thus, the longer one's duration of unemployment the larger is the probability that the individual will be attached to a semi or unskilled job. It could be argued that employers are using duration of unemployment as a productivity signal or screening device in the allocation of higher grade jobs to young workers. The change of residence coefficient fails to show in a well determined way for most of the male categories with the notable exception of the clerical non-manual category. This could be explained by the migration of individuals from outside Dublin into public service clerical jobs.

In terms of the female coefficient estimates of table 6.2 most of the estimates presented are broadly in line with those reported in table 6.1. However, the education coefficient estimates are much more pronounced and well determined for all the occupational categories. The largest effect of any variable in table 6.2 is the the change of residence variable which records a solid positive effect for the female professional category. This may suggest that in order for females to pursue a professional career geographical mobility is necessary. Unemployment appears to play a role in occupational attachment for young females though its effects are not as pronounced as in the male equations and are more concentrated in the lower white collar categories.

The most surprising feature of the female coefficient estimates are those associated with Father's socio-economic background. In general, the coefficients are badly determined in relation to non-manual family background and negative in relation to a manual family background. This suggests that females from an agricultural background fare significantly better than those from a manual family background in terms of white collar occupational attainment. It should be noted that little statistical difference exists between possessing an agricultural or a white collar parental background. The different effects only enter in relation to the manual parental background. A number of tentative interpretations may be provided for this particular finding which was also observed in the structural model estimates of chapter five. One interpretation may lie in the existence of different sets of values across different parental backgrounds. If, for example, the daughters of farmers are encouraged to stay on longer in the education system than those from a manual background due to, perhaps, a higher value being placed on education by farmers, then, this may explain their higher probability of non-manual attachment. Alternatively, farmer's daughters may view urban areas as possessing greater opportunity[58] and since there exists a high correlation between urban areas and non-manual jobs they may find themselves placed in non-manual jobs. On the demand side it may be that young females from rural backgrounds are viewed by non-manual employers as being more reliable than young females from manual backgrounds. In this way they secure non-manual jobs through the prejudices of employers. The truth, however, is probably better represented by a combination of all the above.

In general, the large number of insignificant variables brings in to question the reliability of the predictions. In terms of goodness of fit both occupational models poorly predict the allocation of individuals to their correct occupational groups. The poor predictive power of the occupational models has clear implications for the selectivity bias terms used in the estimation of the ocupational wage equations. The poor estimates of the occupational models could be explained

[58] Since their opportunities in agriculture are limited.

by the small number of observations in some of the occupational categories. Thus, the poor predictive power of the occupational models is the price paid for a finer occupational disaggregation than that of chapter five.

Table 6.1

Reduced Form Male Occupational Estimates

Variable	Prof.& Man.	Clerk.	Other	Skilled
Constant	-3.5865***	-1.2592***	-1.4011***	0.9324***
	(0.4456)	(0.2955)	(0.3406)	(0.2147)
Previous Experience	0.1844**	-0.0870	0.0601	-0.0908**
	(0.0746)	(0.0599)	(0.0588)	(0.0446)
Education	0.9009***	0.4593***	0.1337	0.0572
	(0.0854)	(0.0713)	(0.0876)	(0.0629)
Residence of Schooling				
Dublin City & Co.	0.8791**	1.0645***	0.7732**	0.2352
	(0.4117)	(0.2970)	(0.3355)	(0.2413)
East & Midlands	-0.1499	-0.1274	-0.1173	-0.1398
	(0.2171)	(0.3366)	(0.2954)	(0.4255)
South	0.0899	0.0867	-0.0618	0.1849
	(0.2103)	(0.3292)	(0.2831)	(0.4005)
Unemployment (months)	-0.0856***	-0.0936***	-0.0262**	-0.0578***
	(0.0107)	(0.0135)	(0.0188)	(0.0282)
Change of Residence	0.2114	1.1507**	0.5331	0.5091
	(0.4082)	(0.5379)	(0.4330)	(0.5618)
Father's Occupation				
Non-Manual	1.6954***	1.3191***	1.0076***	0.4365**
	(0.3756)	(0.2804)	(0.3268)	(0.2241)
Manual	0.0504	0.3127	0.3282	0.0257
	(0.3787)	(0.2554)	(0.2947)	(0.1827)

Asymptotic standard errors are in parentheses and * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. The number of observations is 1505.**

Table 6.2

Reduced Form Female Occupational Equations

Variable	Prof.& Man.	Clerk.	Other	Skilled
Constant	-4.4131***	-0.5332*	-0.9495***	-1.8426***
	(0.5477)	(0.2757)	(0.3437)	(0.4631)
Previous Experience	0.2518***	-0.0344	0.1151*	0.0182
	(0.0884)	(0.0553)	(0.0652)	(0.0901)
Education	0.1.1640***	0.8566***	0.4315***	0.3008***
	(0.1115)	(0.0722)	(0.0919)	(0.1142)
Residence of Schooling				
Dublin City & Co.	1.5705***	1.1196***	0.5507	-0.0654
	(0.4737)	(0.2712)	(0.3537)	(0.5120)
East & Midlands	0.9301*	0.5646**	-0.0239	0.2257
	(0.4835)	(0.2715)	(0.3639)	(0.4823)
South	0.8603*	0.6142**	0.5558*	1.1705***
	(0.4515)	(0.2605)	(0.3271)	(0.4213)
Unemployment (months)	-0.0413	-0.0659***	-0.0340*	-0.0129
	(0.0313)	(0.0165)	(0.0202)	(0.0227)
Change of Residence	2.0012***	0.3915	0.9071*	0.0972
	(0.5185)	(0.4622)	(0.5156)	(0.7434)
Father's Occupation				
Non-Manual	0.1601	-0.0399	-0.3953	-0.3734
	(0.4030)	(0.2621)	(0.3262)	(0.4328)
Manual	-1.1775***	-0.5704***	-1.0302***	-0.3040
	(0.4092)	(0.2362)	(0.2992)	(0.3698)

Asymptotic standard errors are in parentheses and * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. The number of observations is 1322.**

6.6 Wage Equation Estimates

Tables 6.3 to 6.6 report the coefficient estimates based on OLS and the consistent estimator outlined in section 6.3. Most of the coefficient estimates appear robust to the estimator used with the exception of the years in post-compulsory education variable. The contrasts between the set of OLS and consistent estimates for this variable is most marked in those equations where there is strong evidence of selectivity bias. Since education appears as a strong determinant of occupational attachment (as the previous section outlined) this need not be interpreted as such a surprising result.

In terms of an economic interpretation the wage equation estimates reveal little in the way of surprise. The returns to labour force experience are steeper in the first years of labour force experience in compont with human capital theory. The returns to an additional year in education though varying across occupational categories are in line with the estimates in chapters four and five. Nevertheless, the variability in the female returns is much greater and in the skilled occupational category the returns to formal post-compulsory education are not statistically different from zero. For most of the male occupational categories residing in Dublin has a pronounced positive effect and also wages appear to increase significantly with increases in firm size. The same could in general be said for the female estimates though the incidence of significance on the Dublin coefficients are less. One other contrasting feature worthy of mention relates to the unemployment variable. For all the male occupational categories a positive relationship exists between unemployment and the wage. This is a finding not inconsistent with the predictions of neoclassical job search theory. However, for the females a negative relationship is recorded for all three white collar occupational categories which suggests a differential treatment of females in regard to unemployment and wages. However, in terms of the consistent estimates only the clerical occupational category records a statistically significant effect.

Attention now turns to the selectivity bias terms and their interpretation. Though the asymptotic properties of the t-statistic associated with the $-\frac{\phi}{\Phi}$ term are not well known[59] they

[59] This is in contrast to the Heckman procedure where Melino (1982) provides an LM test for exogeneity which was used in chapter five.

may be used as a rough guide to ascertain if there exists evidence of non-randomness in the allocation of workers to occupational categories. The signs on the estimated proxy variables also provide information on the direction of the selectivity bias. Due to the construction of the proxy variables a positive (negative) coefficient implies negative (positive) selectivity or truncation in terms of the occupational category in question. The economic interpretation for this is that a worker with a given set of characteristics selected at random from the population faces on average a lower (higher) wage offer distribution if selected into that sector rather than any other sector.

For the male equations the strongest selectivity effect emerges for the skilled occupational category and suggests that for a young male worker drawn at random from the population with average characteristics the observed wage offer distribution is 15% less than the distribution that would be observed for the average male individual selected into any other occupational category. Human capital theory can again be invoked to provide an explanation for this particular finding. It suggests that in the context of firm specific human capital training young workers accept a wage less than their marginal product in order to share the costs of training with employers and this will be particularly so for young skilled workers.

An even stronger selectivity effect in the female Skilled sector is recorded suggesting that on average the net hourly wage distribution is almost 30% less than the wage distribution observed for the same individual selected into any other sector. However, the small number of observations in this particular category suggests a cautious interpretation as the coefficients are not well determined for a number of key variables. The other strong selectivity effect recorded for the females is in terms of the clerical occupational sector. The coefficient here suggests that the wage offer distribution that is observed for the average female drawn at random from the population into this category is about 16% above what would be observed for the average female with the same personal characteristics selected into any other occupational category. Neither of these results should come as any surprise. Young females who end up in secure white collar clerical jobs tend on average to do better than their peers. This is certainly the case in the context of young workers. It is clear that both these findings and that obtained for the skilled male category vindicate the use of the consistent estimator.

A tentative test for the joint statistical significance of the selectivity bias terms and hence occupational exogeneity is provided by a likelihood ratio test. The χ^2 statistics with five degrees of freedom[60] for the male and female equations are 7.00 and 7.37 respectively suggesting non-rejection of the null hypothesis of exogeneity. Nevertheless, the rejection of the joint significance of the selectivity terms conceals the individual effects which emerge strongly in the skilled occupational categories for both sexes and the clerical category for the females.

[60] The associated critical values are 15.09, 11.07 and 9.24 for the 1%, 5% and 10% level of significance.

Table 6.3

Male Occupational Wage Estimates (Consistent)

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof.& Man
Constant	-0.0983	-0.4715***	-0.1054	0.0559	-0.2118
	(0.1091)	(0.0504)	(0.1158)	(0.1898)	(0.2128)
Exp. ≤ 4 yrs.	0.1077***	0.1708***	0.0744**	0.1035***	0.1104***
	(0.0218)	(0.0120)	(0.0291)	(0.0143)	(0.0287)
Exp.> 4 yrs.	0.0351**	0.0604***	0.0550**	0.0466***	0.0556
	(0.0146)	(0.0098)	(0.0223)	(0.0181)	(0.0444)
Education	0.0794***	0.0459***	0.0728***	0.0321	0.0759***
	(0.0258)	(0.0133)	(0.0215)	(0.0204)	(0.0281)
50 ≤ Firm <400	0.1565**	0.1127***	0.1925***	0.1274**	0.0672
	(0.0587)	(0.0286)	(0.0533)	(0.0555)	(0.0688)
Firm ≥ 400	0.1977***	0.1154***	0.2930***	0.2095***	0.0794
	(0.0426)	(0.0309)	(0.0589)	(0.0429)	(0.0789)
Current Residence					
Dublin City & Co.	0.1236*	0.0617*	0.1086	0.0728	0.2591**
	(0.0624)	(0.0371)	(0.0792)	(0.0700)	(0.1038)
South	0.0739	0.0992***	0.1132	-0.0526	0.1426
	(0.0649)	(0.0338)	(0.0764)	(0.0712)	(0.1049)
Midlands	-0.0336	-0.0205	-0.0363	-0.0179	0.0409
	(0.0571)	(0.0356)	(0.1305)	(0.0656)	(0.1626)
East	0.0240	0.1185***	0.2268***	-0.0344	0.0213
	(0.0936)	(0.0453)	(0.0803)	(0.0696)	(0.1320)
Unemployment (months)	0.0040*	0.0023	0.0038	0.0036	0.0103
	(0.0024)	(0.0028)	(0.0033)	(0.0043)	(0.0091)
Change of Residence	0.1432**	0.0553	0.1481	0.0362	0.0967
	(0.0577)	(0.0525)	(0.1157)	(0.0454)	(0.1127)
-[illegible]	-0.0848	0.1515***	-0.0544	-0.0729*	-0.0070
	(0.0668)	(0.0579)	(0.0286)	(0.0935)	(0.0637)
Observations	277	660	146	296	126

Standard errors are in parenthesis and are based on White (1980) and are heteroscedastic consistent. * denotes statistical significance at the 1% level, ** at the 5% level and * at the 10% level using two tailed tests.**

Table 6.4

Male Occupational Wage Estimates (OLS)

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof.& Man.
Constant	-0.2049***	-0.3891***	-0.1921	-0.0757	-0.2347*
	(0.0664)	(0.0419)	(0.1188)	(0.0817)	(0.1386)
Exp. ≤ 4 yrs.	0.1111***	0.1714***	0.0740**	0.1035***	0.1112***
	(0.0204)	(0.0120)	(0.0292)	(0.0143)	(0.0294)
Exp.> 4 yrs.	0.0344**	0.0641***	0.0528**	0.0444**	0.0555
	(0.0145)	(0.0096)	(0.0223)	(0.0173)	(0.0447)
Education	0.0644***	0.0639***	0.0605***	0.0401**	0.0783***
	(0.0174)	(0.0104)	(0.0180)	(0.0171)	(0.0187)
10 ≤ Firm <400	0.1564***	0.1171***	0.1937***	0.1292**	0.0672
	(0.0585)	(0.0285)	(0.0534)	(0.0558)	(0.0687)
Firm ≥ 400	0.1956*	0.1136***	0.3019***	0.2154***	0.0798
	(0.0427)	(0.0310)	(0.0584)	(0.0424)	(0.0786)
Current Residence					
Dublin City & Co.	0.0884	0.1030***	0.1186	0.1047*	0.2619
	(0.0523)	(0.0330)	(0.0797)	(0.0576)	(0.1061)
South	0.0614	0.0943***	0.1042	-0.0450	0.1443
	(0.0614)	(0.0339)	(0.0762)	(0.0692)	(0.1078)
Midlands	-0.0306	-0.0137	-0.0496	-0.0143	0.0406
	(0.0573)	(0.0353)	(0.1310)	(0.0664)	(0.1618)
East	0.0256	0.1303***	0.2274***	-0.0272	0.0195
	(0.0940)	(0.0450)	(0.0800)	(0.0699)	(0.1287)
Unemployment (months)	0.0060***	0.0041	0.0047	0.0019	0.0103
	(0.0016)	(0.0028)	(0.0032)	(0.0043)	(0.0088)
Change of Residence	0.0953**	0.0546	0.1080	0.0561	0.1146
	(0.0552)	(0.0524)	(0.1090)	(0.0417)	(0.1125)
Observations	277	660	146	296	126

Standard errors are in parenthesis and are based on White (1980) and are heteroscedastic consistent. * denotes statistical significance at the 1% level, ** at the 5% level and * at the 10% level using two tailed tests.**

Table 6.5

Female Occupational Wage Estimates (Consistent)

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof & Man
Constant	0.00173	-0.6847*	-0.7640***	-0.0057	-0.2304
	(0.0649)	(0.4061)	(0.1679)	(0.0852)	(0.3191)
Exp. ≤ 4 yrs.	0.0557***	0.0598	0.1758***	0.0721***	0.0833
	(0.0123)	(0.0463)	(0.0264)	(0.0084)	(0.0476)
Exp.> 4 yrs.	0.0054	0.0492**	0.0285	0.0442***	0.0931
	(0.0099)	(0.0222)	(0.0249)	(0.0121)	(0.0342)
Education	0.0466	-0.0225	0.0956***	0.0444***	0.1070
	(0.0286)	(0.0557)	(0.0241)	(0.0140)	(0.0369)
50 ≤ Firm <400	0.1161***	0.1523	0.3075***	0.1691***	-0.0501
	(0.0340)	(0.1385)	(0.0693)	(0.0244)	(0.0987)
Firm ≥ 400	0.2480***	0.4150***	0.3129***	0.2419***	0.2040**
	(0.0369)	(0.1186)	(0.0767)	(0.0205)	(0.0915)
Current Residence					
Dublin City & Co.	0.0597	-0.0260	0.1003	0.1813***	0.1061
	(0.0385)	(0.1673)	(0.0849)	(0.0333)	(0.0685)
South	0.0787	0.1601**	0.0355	0.0927***	0.1316
	(0.0382)	(0.0988)	(0.0750)	(0.0341)	(0.0854)
Midlands	0.0251	-0.0636	0.0585	0.0117	0.4988*
	(0.0432)	(0.1863)	(0.1005)	(0.0407)	(0.2509)
East	-0.0472	-0.0633	0.3931**	0.0582	0.3077
	(0.0509)	(0.1443)	(0.1541)	(0.0454)	(0.1887)
Unemployment (months)	0.0056**	0.0122**	-0.0012	-0.0043	-0.0050
	(0.0024)	(0.0056)	(0.0049)	(0.0029)	(0.0107)
Change of Residence	0.1223*	0.0788	0.0450	0.0152	0.0053
	(0.0658)	(0.0787)	(0.0960)	(0.0441)	(0.1015)
$-\frac{\phi}{\Phi}$	-0.0321	0.2800*	0.0337	-0.1605**	0.0092
	(0.0659)	(0.1615)	(0.0959)	(0.0776)	(0.0478)
Observations	251	73	145	771	82

Standard errors are in parenthesis and are based on White (1980) and are heteroscedastic consistent. * denotes statistical significance at the 1% level, ** at the 5% level and * at the 10% level using two tailed tests.**

Table 6.6

Female Occupational Wage Estimates (OLS)

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof.& Man
Constant	-0.0185	-0.1289	-0.7120***	-0.1614***	-0.1942
	(0.0491)	(0.1445)	(0.0991)	(0.0407)	(0.2464)
Exp. ≤ 4 yrs.	0.0556***	0.0633	0.1741***	0.0704***	0.0816*
	(0.0123)	(0.0485)	(0.0259)	(0.0084)	(0.0469)
Exp.> 4 yrs.	0.0053	0.0509**	0.0274	0.0408***	0.0962***
	(0.0099)	(0.0238)	(0.0248)	(0.0119)	(0.0371)
Education	0.0341***	0.0439*	0.1011***	0.0676***	0.1038***
	(0.0111)	(0.0263)	(0.0181)	(0.0072)	(0.0288)
50 ≤ Firm <400	0.1145***	0.1476	0.3063***	0.1738***	-0.0507
	(0.0337)	(0.1403)	(0.0697)	(0.0240)	(0.0992)
Firm ≥ 400	0.2480***	0.3752***	0.3110***	0.2457***	0.2038**
	(0.0370)	(0.1070)	(0.0767)	(0.0205)	(0.0917)
[illegible]					
Dublin City & Co.	0.0487	0.1375	0.1081	0.2108***	0.1017
	(0.0310)	(0.1040)	(0.0781)	(0.0306)	(0.0651)
South	0.0716**	0.0197	0.0376	0.0986***	0.1303
	(0.0344)	(0.1256)	(0.0744)	(0.0342)	(0.0846)
Midlands	0.0210	-0.0476	0.0667	0.0291	0.5018**
	(0.0414)	(0.1872)	(0.0949)	(0.0399)	(0.2397)
East	-0.0503	-0.0671	0.4030***	0.0777*	0.3067
	(0.0508)	(0.1374)	(0.1495)	(0.0453)	(0.1898)
Unemployment (months)	0.0062***	0.0096**	-0.0012	-0.0073***	-0.0051
	(0.0021)	(0.0048)	(0.0048)	(0.0025)	(0.0107)
Change of Residence	0.1068*	0.1418*	0.0330	-0.0268	-0.0013
	(0.0622)	(0.0655)	(0.0943)	(0.0368)	(0.0964)
Observations	251	73	145	771	82

Standard errors are in parenthesis and are based on White (1980) and are heteroscedastic consistent. * denotes statistical significance at the 1% level, ** at the 5% level and * at the 10% level using two tailed tests.**

6.7 Occupational Wage Differentials

The mean observed gender wage differential across occupations is (in logarithms) 0.0282. In other words, males on average earn just under 3% more in wages than do females. The mean gender wage offer differential across occupations is less suggesting a negligible aggregate differential at 0.0012. These relatively low aggregate figures mask vast differences within given occupations as is evident from tables 6.7 and 6.8. The observed wage differentials range from close to zero for the professional and managerial category to over 30% for the other intermediate non-manual category. The gulf in magnitude is even more dramatic in terms of the occupational wage offer differentials. These vary from 2.8% in the professional and managerial category to over 50% in the skilled sector. As is recalled from chapter five focusing exclusively on the observed wage differences within given occupations obscures important effects if the female equations are characterised by a strong selectivity effect. This is seen in terms of Table 6.8 where the widest wage offer differentials exist, not surprisingly, in those occupational categories where there is either strong positive selectivity effects in the male equations and/or strong negative selectivity effects in the female equations.

Using the wage equation estimates of tables 6.3 to 6.6 observed wage and wage offer differentials by occupation can be decomposed into explained and unexplained parts. As Cain (1986) surveys both parts of the differential can be viewed as providing a handle on discrimination. The explained part of the differential may be interpreted as representing the effects of pre-entry discrimination if females are assumed to encounter institution-related discrimination in their procurement of productivity enhancing characteristics. For example, if females encounter discrimination in their subject uptake at the secondary education level than the explained part of the differential may capture the pre-entry effect. The unexplained part is interpreted as the post-entry labour market discrimination effect and most attention focuses on this particular component.

As Table 6.7 highlights the unexplained differentials in observed wages record vast gulfs in magnitude with the largest and most significant effect obtained for the other intermediate non-manual category. However, when selectivity bias is accounted for the unexplained wage offer differential in this category increases from 21% to over 38%. Nevertheless, the most resounding increase is reserved for the skilled occupational category where the unexplained differential increases to over 50%. The relatively small number of observations in both the female skilled sector and the intermediate non-manual sector suggests caution in interpretation. Nevertheless, the results illustrate the importance of controlling for selectivity bias in occupational wage equations. Though tables 6.7 and 6.8 are dealing with different concepts it is clear that within given occupations the female disadvantage, however measured, is severe. This is particularly the case in the two occupational categories highlighted above. It could be argued that since only 16% of all females are in both of these occupational categories together, there need be little cause for concern. This is not a view that should be accepted too lightly, however. Accepting that both males and females are sacrificing wages, in the skilled sector, in order to undergo training it's clear that the relative sacrifice is falling disproportionately on the females. As is clear from the selectivity bias coefficients and the wage offer differentials in the skilled category the sacrifice the female workers must make in wages in order to undergo training is nearly twice that for males. This has clear implications for females in terms of their willingness to undergo training and has long-term implications for their occupational segregation. It cannot be stated from the above that the vast wage offer differential relates to the type of skilled jobs females choose or to the inordinately higher cost imposed on females by employers in comparison to males. Whatever the explanation it is clear from the analysis that the incentives for females to undergo training in the skilled sector are significantly less than for males.

It is also clear from both tables 6.7 and 6.8 that some of the differences in wages are explained by differing endowments. If discrimination exists in terms of access to these endowments these estimates could be interpreted as reflecting the effects of some form of pre-entry discrimination. The strongest positive effects for this component of the wage decomposition come through in terms of the clerical and the other intermediate non-manual occupational category. It could be argued that if males in both these categories have greater access to the qualifications necessary for securing promotion than females then these components could be interpreted as discrimination estimates. However, it cannot be unambiguously stated that this part of the differential reflects discrimination. Nevertheless, failure to detect statistically significant unexplained differentials should not be interpreted as absence of explicit wage discrimination in given occupations.

Table 6.7

Gender Wage Differentials by Occupation

Occupation	$\Delta\bar{W}_i$	$\Delta\bar{Z}_i\beta^m$	$\bar{Z}_i\Delta\beta$
Prof. & Man.	-0.0014	-0.0275	0.0261
		(0.0317)	(0.0523)
Clerical	0.0586	0.0402***	0.0184
		(0.0080)	(0.0219)
Other White	0.2641	0.0520***	0.2121***
		(0.0167)	(0.0415)
Skilled	0.0190	-0.0421***	0.0611
		(0.0108)	(0.0414)
Semi & Unskilled	0.0196	-0.0179	0.0375
		(0.0163)	(0.0307)

Asymptotic standard errors in parenthesis. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests.**

Table 6.8

Gender Wage Offer Differentials by Occupation

Occupation	$\Delta\bar{W}_i^c$	$\Delta\bar{Z}_i\beta^m$	$\bar{Z}_i\Delta\beta$
Prof. & Man.	0.0280	-0.0285	0.0565
		(0.0326)	(0.1624)
Clerical	0.0592	0.0387***	0.0205
		(0.0080)	(0.1312)
Other White	0.4253	0.0426***	0.3827*
		(0.0187)	(0.1845)
Skilled	0.4931	-0.0433***	0.5364
		(0.0108)	(0.3705)
Semi & Unskilled	0.0969	-0.0223	0.1192
		(0.0180)	(0.1383)

Asymptotic standard errors in parenthesis. * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. c denotes corrected for selectivity bias.**

6.8 Occupational Distributions

Table 6.9 provides actual and simulated occupational distributions for males and females across the five occupational categories considered. Column 1 gives the male means, column 2 the female means and column 3 the female means if a male occupational structure is imposed. Column 3 is obtained by fitting the male occupational coefficients of table 6.1 to the female realisations of the explanatory variables. The most obvious feature of the table is the dramatic shift of females away from the clerical occupational category to the male dominated skilled occupational sector. Furthermore, in the context of the above outlined scenario the female representation would rise from 6.2% to just under 10% in the professional and managerial category and the proportion of females in the semi and unskilled category would fall from almost 20% to just over 15%. The female representation within the other intermediate non-manual occupational category remains relatively stable with the imposition of the male occupational structure. It is clear from the above evidence that females are over-represented in the low skill, low pay sectors and under-represented in the skilled and professional sectors

An index of occupational dissimilarity proposed by Duncan and Duncan (1955) is used to provide some objective measure of occupational differences between young males and females. The index is defined, in general, as

$$\mathbf{D} = (1/2)\sum_{j=1}^{M} \mid \bar{\mathbf{P}}_j^m - \bar{\mathbf{P}}_j^f \mid \tag{6.18}$$

and in this case gives the proportion of a group, male and/or female, that would have to shift to equalise the gender representation across occupations. A variant of this index is also calculated based on the female predicted outcomes in column three of table 6.9. This index may be given by

$$\mathbf{D}^* = (1/2)\sum_{j=1}^{M} \mid \bar{\mathbf{P}}_j^m - \hat{\mathbf{P}}_j^f \mid . \tag{6.19}$$

As illustrated in table 6.9 the degree of dissimilarity suggested by the D index is reasonably high. It suggests that 40.5% of either or a combination of both sexes would have to shift to equalise sexes across occupational categories. Not surprisingly, when D* is calculated only 4.4% of either sex would have to shift to effect gender equalisation. This suggests that only 11% of the observed disparity in occupational attachment is explained by characteristics or endowments with over 89% interpretable as demand side discrimination, supply side preferences or some indeterminate combination of both.

Using tables 6.7, 6.8 and 6.9 the mean wage differential of 0.0282 can be decomposed into inter and intra-occupational differences in wages. The intra-occupational effect is estimated as 0.0678 and the inter-occupational effect as -0.0397. The sum of the two unjustified components, wage discrimination and occupational segregation[61], is 0.0121 which represents nearly 43% of the gross observed gender difference in wages. Thus, just over 1% of the average net hourly male wage may be interpreted as an occupational/wage discrimination effect. This by any standards is small and though no standard errors are estimated for this effect its magnitude may not be deemed significant in an economic sense. The same exercise is performed for the consistent wage equation estimates with the results recorded in the second column of table 6.10. Though the magnitudes are slightly different the same story holds. The greatest part of the female wage disadvantage lies not in their occupational distribution but in their wage disadvantage within given occupations, for example, the other intermediate non-manual and the skilled occupational category

The significant finding of the above analysis is the effect of the unexplained intra-occupational wage differences on the overall mean wage differential. This may best be illustrated by table 6.11. In addition to the mean male and female wages table 6.11 contains two predicted female wages based on the OLS and the consistent estimator. The first predicted wage (column three) simply allows female access to male occupations assuming the given female wage structure in these occupations. It is clear that allowing females access to male occupations has no significant impact on their relative wage position. In the OLS case cited here it actually leads to a decrease in their wage position. This is explained by the fact that the imposed male occupational structure allocates a large proportion of females to the lower paid skilled category. A similar finding and a more dramatic one is found for the consistent estimator. The relatively low predicted wage is explained here by the strong negative selectivity effect in this sector and the consequent low wage offer that confronts females in this male-intensive sector.

The second predicted female wage (column four) is obtained by imposing both a male occupational structure and a male wage structure and leads to a relatively dramatic rise in the female mean wage for both the OLS and the consistent estimates. This is in comport with the

---

[61] Occupational segregation is interpreted as unjustified in this particular case since it refers to the female occupational distribution that is not justified, relative to the male occupational distribution, on the basis of the female characteristics.

findings of both Brown *et al.* (1980) and Miller (1987) and prompted the conclusion on both their parts for a more adequate anti-discrimination legislation aimed at promoting equal pay within occupations rather than aimed at gender equalisation across occupations.

Though the magnitude of the mean wage differences obtained in this chapter is relatively small in comparison to the two studies cited above, this may be explained in large part by both the single status and the youth of the workers considered. Nevertheless, one clear conclusion emerges. The disadvantaged female wage position is not explained to any great degree by their occupational distribution. The greater weight of evidence appears to suggest that the disadvantage operates within the (albeit broadly defined) occupational groups and remedies aimed at ameliorating the promotional prospects of women as well as their training prospects within given occupations may be a more effective policy than one designed at gender equalisation.

Table 6.9

Occupational Distributions and Dissimilarity Indices.

Occupation	$\bar{P}_j^m$	$\bar{P}_j^f$	$\hat{P}_j^f$	$\lvert \bar{P}_j^m - \bar{P}_j^f \rvert$	$\lvert \bar{P}_j^m - \hat{P}_j^f \rvert$
Prof. & Man.	0.0837	0.0620	0.0956	0.0217	0.0119
Clerical	0.1967	0.5832	0.2293	0.3865	0.0326
Other White	0.0970	0.1097	0.0929	0.0127	0.0041
Skilled	0.4385	0.0552	0.4252	0.3833	0.0133
Semi & Unskilled	0.1841	0.1899	0.1569	0.0058	0.0272
D	-	-	-	0.4050	-
D*	-	-	-	-	0.0445

Table 6.10

Wage Decomposition

Component	Wage [a]	Wage [b]	Description
ΔW	0.0282	0.0012	Aggregate Difference
$\sum_{j=1}^{k} \bar{P}_j^f \, \Delta\bar{Z}_j \, \hat{\beta}_j^m$	0.0217	0.0189	Differing Characteristics
$\sum_{j=1}^{k} \bar{P}_j^f \, \bar{Z}_j^f \, \Delta\hat{\beta}$	0.0462	0.1097	Differing Parameters
$\sum_{j=1}^{k} \overline{W}_j^m (\bar{P}_j^m - \hat{P}_j^f)$	-0.0062	-0.0075	Differing Characteristics
$\sum_{j=1}^{k} \overline{W}_j^m (\hat{P}_j^f - \bar{P}_j^f)$	-0.0335	-0.1199	Differing Parameters

a denotes estimates based on OLS procedure of Tables 6.4 and 6.6. *b* denotes estimates based on the consistent procedure of Tables 6.3 and 6.5.

Table 6.11

Simulated Wage Comparison

Wage	$\sum_{j=1}^{5}\bar{P}_j^m \bar{Z}_j^m \beta_j^m$	$\sum_{j=1}^{5}\bar{P}_j^f \bar{Z}_j^f \beta_j^f$	$\sum_{j=1}^{5}\bar{P}_j^{f'} \bar{Z}_j^f \beta_j^f$	$\sum_{j=1}^{5}\bar{P}_j^{f'} \bar{Z}_j^f \beta_j^m$
OLS	0.4125	0.3843	0.3696	0.4279
Consistent	0.4069	0.4057	0.1338	0.4262

6.9 Conclusions

The main objectives of this chapter have been three-fold. Firstly, to provide estimates for some of the key determinants of occupational attachment for young workers in Ireland. Secondly, to use this information to correct for the potential existence of selectivity bias in occupational wage equations and thirdly to quantify the gender differential in wages that operates through occupational segregation. In examining this last issue a distinction is made between observed differences in wages and differences in wage offers (*i.e.* observed wages corrected for selectivity bias).

The first issue was addressed with the estimation of the unordered multinomial logit using a five-way occupational categorisation. Education, residence of schooling, unemployment and parental background all featured as prominent determinants of occupational attachment for both males and females. In terms of one's attachment to the non-manual occupational categories education and schooling in Dublin were both found to play an important role. Unemployment played a stronger negative role in occupational attachment for males than for females. However, the biggest contrast in results occurred in terms of the parental background variable. A non-manual background is more likely to see a young male in a non-manual job with the reverse almost being the case for a young female.

The second issue addressed was that of selectivity bias in the occupational wage equations. Contrasting the OLS estimates with those from a consistent estimator revealed evidence of selectivity bias which was strongest in the skilled occupational category for both males and females. This had important implications for the discrimination estimates in the skilled category where the unexplained differential was found to contrast dramatically across the two estimators. It also suggested the existence of a differential in incentives available to males and females in terms of training in the skilled sector. This is seen has having long-term implications for the occupational segregation of females.

The third issue was the decomposition of the mean wage differential (in both its forms) into its component parts. Neither mean differential was found to be large and the low mean effects were found to mask dramatic occupational wage differentials. The widest differentials were reported for the consistent occupational differentials and in particular for those categories which exhibited evidence of selectivity bias. In line with Miller (1987) and Brown *et al.* (1980) little of either the aggregate wage offer differential or observed wage differential was found to be due to occupational segregation. In the vernacular of Sloane (1985) partial segregation by occupation appears to explain a greater part of the aggregate wage differential. In this case the female wage disadvantage operates through their segregation within given occupations rather than across occupations. Policies designed at equalising gender representation across occupations may, therefore, be less effective than policies aimed at dismantling their partially segregated positions within given occupations. Increased access to training and promotion in such a way as to extend the career ladders confronting females may be a more potent measure in removing their disadvantaged status.

Chapter Seven

Wage Differentials and the Dual Labour Market

7.1 Introduction

As chapter two surveyed neo-classical theory attempts to explain the existence of wage differentials by recourse to, *inter alia*, human capital theory. In this framework differentials in wages are explained by differentials in the accumulation of human capital investment by the individual. Labour market segmentation theory, on the other hand, attempts to explain wage differences on the basis of the type of job an individual holds. A major limitation of the labour market segmentation literature lies in the fact that it does not provide a single or unique explanation for the existence of wage differentials. This in some sense is reflected in the empirical work that has hitherto been characterised by a relatively *ad hoc* approach.

To test one of the basic premises of segmented and dual labour market (SDLM) theory, that of the existence of two distinct wage-setting mechanisms, researchers invariably allocated workers to primary/secondary jobs on the basis of their job characteristics and/or their industry and/or occupation depending on the segment of interest. A statistical test of the difference between wage coefficient estimates provided the basic test. An approach similar to this is adopted by Osterman (1975), Carnoy and Rumberger (1980) and Zucker and Rosenstein (1981) in their empirical studies. Conflicting findings emerge from these studies. The conflict in results may be related to an inaccurate classification of jobs. Since invariably the sample of workers is divided into a high wage or a low wage sector the sample is inevitably truncated and the potential problem of selectivity bias emerges. However, in all the above studies this problem is conveniently overlooked. Heckman and Hotz (1986) provide an application in a third world context where poverty status is used as an allocating mechanism and selectivity bias is controlled. They, however, conclude that the dual labour market hypothesis provides insufficient structure to distinguish it from the market-clearing hypothesis. In view of this, they argue, the hypothesis of dualism is untestable.

Other studies have attempted to focus on another tenet of the dual labour market theory, that of inter-sectoral immobility. If primary jobs are rationed as dual labour market theorists suggest then upward mobility would be impossible. Leigh (1976) and Schiller (1977) both provide evidence of upward mobility of individuals and argue this as a refutation of dualism. It could be argued that immobility is consistent with human capital theory predictions (firm specific investments) while mobility is not necessarily inconsistent with the existence of career ladders and internal labour markets within the primary sector. Thus, testing for mobility sheds little light on the dual labour market hypothesis.

More recently Dickens and Lang (1985, 1987 and 1988) outline an alternative econometric approach that incorporates a formal test of the rationing hypothesis. The econometric approach involves the estimation of a switching model with unknown regimes. The full model consists of two wage equations and an equation determining sectoral attachment. By using the standard stochastic switching model (see Maddala (1983), pp. 283-286) no information on the sectoral attachment of the individual is required. Thus no *ad hoc* allocation of workers to sectors need be made. However, the statistical test outlined by Dickens and Lang (1985) to test for the existence of non-price barriers has been the subject of some criticism. Heckman and Hotz (1986) argue that the rejection of the null hypothesis of no dualism by Dickens and Lang is uninformative since the rejection may be more related to departures from the strong assumptions made by Dickens and Lang.[62] Moreover, Heckman and Hotz (1986) argue that the likelihood ratio test constructed by Dickens and Lang is in error since under the null of no dualism the sectoral attachment probabilities are not defined[63]. Though the view taken by Heckman and Hotz is that dualism is not a

[62] Heckman and Hotz argue that the rejection may occur as a consequence of,
(i) more than two sectors in the labour market,
(ii) agents may be utility maximisers rather than income maximisers,
(iii) inappropriate distributional assumptions,
(iv) costs of mobility across sectors.

[63] Use of the Wald test is inappropriate since the coefficients on the switching equation are not identified.

testable proposition the switching regression approach outlined by Dickens and Lang does provide an opportunity to assess at the very least the existence of two distinct wage setting mechanisms which are consistent with the dual labour market description.

Neither of the two broad approaches outlined above could be deemed totally satisfactory. The *ad hoc* approach allocates individuals by brute force on the basis of an investigator's priors with the consequent results invariably providing a self-fulfilling prophecy. The Dickens and Lang approach, on the other hand, places the allocation in the hands of a stochastic switching regression model. Attempting to improve on both approaches provides one of the two motivations for this chapter. The alternative approach proposed merges aspects of both the previous approaches. It is clear that some jobs are unambiguously primary in character and others secondary. There are, however, another sub-set of jobs for which a high degree of ambiguity attaches and hence creating an inability to allocate unequivocally to either sector. In the following section a maximum likelihood approach is outlined based on Stewart (1988) that allows for partial observability in the latent dependent variable and possesses the stochastic switching model employed by Dickens and Lang (1985) as a special case. The framework to be more extensively outlined below allows certain jobs to be allocated to either the primary or the secondary sector with probability one with the residual allocated on the basis of the stochastic switching model.

The second motivation, and a more important one, is to examine the evidence for the existence of a dual labour market for young workers in Ireland. The young workers labour market may be viewed by some as part of a labour market secondary to an adult-dominated primary market. Nevertheless, it would be difficult to dispute that within the young workers labour market there co-exists "good" and "bad" jobs. The proportion of these jobs may be significantly different in the adult and young workers markets but it is clearly of interest to see if the facts are in comport with the dual labour market paradigm as broadly defined. To assess the dual labour market view inter-sectoral wage differentials are calculated under a number of alternative hypotheses and a very informal test of rationing is also provided to assess the degree of sectoral

A normalisation would be required but the Wald test is not invariant to the normalisation used.

immobility.

The layout of the chapter is as follows: the following section outlines the econometric methodology and section 7.3 describes the data and the priors used in the allocation of jobs to sectors. Section 7.4 presents results based on the econometric model outlined in section 7.2 and section 7.5 presents estimates of the wage gap. Section 7.6 provides the informal test of rationing and section 7.7 concludes.

7.2 Econometric Methodology

Following Dickens and Lang (1985) the system describing sectoral wages and sectoral attachment may be expressed in terms of the following set of equations:

$$\mathbf{W}_i = \mathbf{X}_i \beta_p + \varepsilon_{pi} \tag{7.1}$$

$$\mathbf{W}_i = \mathbf{X}_i \beta_s + \varepsilon_{si} \tag{7.2}$$

$$y^* = \mathbf{Z}_i \gamma + \varepsilon_{wi} \tag{7.3}$$

where

p denotes primary and s secondary sector,

$\mathbf{W}_i$ is the log of the net hourly wage,

$\mathbf{X}_i$ is a vector of characteristics determining the wage,

$\mathbf{Z}_i$ is a vector of characteristics determining sectoral attachment,

β_p, β_s, γ are vectors of unknown parameters

ε_p, ε_s, ε_w are error terms and

y_i^* is a latent dependent variable measuring sectoral attachment.

The vector of error terms are assumed to possess a trivariate normal distribution and are distributed with a mean of 0 and a variance given by the 3x3 matrix Ω. Since the errors cannot validly be assumed independent the Ω matrix is expressed as:

$$\Omega = \begin{vmatrix} \sigma_p^2 & \sigma_{ps} & \rho_p\sigma_p \\ \sigma_{sp} & \sigma_s^2 & \rho_s\sigma_s \\ \rho_p\sigma_p & \rho_s\sigma_s & 1 \end{vmatrix}$$

where $\rho_p\sigma_p$ and $\rho_s\sigma_s$ are the covariances between the error in the switching equation and the wage equations of (7.1) and (7.2) respectively. σ_{jl} is the error covariance of the j^{th} equation where $j, l = p, s$. Since the wage equations are not observed simultaneously σ_{ps} is not estimable. As Stewart (1988) points out since the scale and origin of y^* is arbitrary a rescaling is effected such that the variance of e_w is normalised to unity. However, it is not possible to simultaneously identify all the switching equation parameter estimates and its error variance[64].

Though y_i^* is not observed it is assumed that if $y^* > 0$ the individual's wage is determined by (7.1) and if $y^* \leq 0$ it is determined by (7.2)[65]. Thus if $y^* > 0$ this implies

$$e_{wi} > -Z_i\gamma$$

and if $y^* \leq 0$, then

$$e_{wi} \leq -Z_i\gamma$$

The conditional density for the switching equation's error is given by

$$e_{wi} \mid e_j \quad N\left(\frac{e_j\rho_j}{\sigma_j},\ 1 - \rho_j^2\right)$$

and the required probabilities are given by

$$prob[e_{wi} > k \mid e_j] = 1 - \Phi\frac{\left[k - \frac{e_j\rho_j}{\sigma_j}\right]}{\sqrt{1 - \rho_j^2}}$$

where in both cases $j = p, s$

The likelihood for the problem can now be expressed as

$$prob\,(e_{wi} > -Z_i\gamma \mid Z_i, X_i, e_{pi}).\ f(e_{pi}) + prob\,(e_{wi} \leq -Z_i\gamma \mid Z_i, X_i, e_{si}).\ f(e_{si}) \qquad (7.4)$$

[64] The exogenous switching model emerges if $\rho_s\sigma_s = \rho_p\sigma_p = 0$.

[65] Economic theory and the signs of the γ vector of coefficients must be used ultimately to determine which sector the attachment equation is predicting.

where everything is as above and f( ) is the marginal density function of the associated error term. The likelihood for the i^{th} individual may now be written as follows

$$L_i = \left[1 - \Phi\left[\frac{-Z_i\gamma - \frac{\rho_p}{\sigma_p}(W_i - X_i\beta_p)}{\sqrt{1-\rho_p^2}}\right]\right]\frac{1}{\sigma_p}\phi\left[\frac{W_i - X_i\beta_p}{\sigma_p}\right] \quad (7.5)$$

$$+\Phi\left[\frac{-Z_i\gamma - \frac{\rho_s}{\sigma_s}(W_i - X_i\beta_s)}{\sqrt{1-\rho_s^2}}\right]\frac{1}{\sigma_s}\phi\left[\frac{W_i - X_i\beta_s}{\sigma_s}\right]$$

where $\Phi(\ )$ and $\phi(\ )$ are the cumulative normal distribution and density functions respectively.

For convenience denote

$K_{ji} = \dfrac{W_i - X_i\beta_j}{\sigma_j}$ and

$V_{ji} = \dfrac{Z_i\gamma + \rho_j K_{ji}}{\sqrt{1-\rho_j^2}}$.

Now the log likelihood may be written more compactly as:

$$L = \sum_{i=1}^{T} \log\left[\Phi(V_{pi})\frac{\phi(K_{pi})}{\sigma_p} + (1 - \Phi(V_{si}))\frac{\phi(K_{si})}{\sigma_s}\right] \quad (7.6)$$

where L is the log likelihood and T is the number of observations and everything else is as defined above.

Maximum likelihood techniques are used in estimation and on the basis of the γ vector of estimates unconditional (*ex ante*) probabilities and conditional (*ex post*) probabilities of sectoral attachment are obtained. The unconditional probability for the i^{th} individual is denoted by $\Phi(Z_i\gamma)$ and Bayes theorem is used to obtain the conditional probability for the primary sector (conditional on the sectoral wage) for the i^{th} individual. In terms of this analysis the conditional probabilities are given by

$$\frac{\Phi(V_{pi})\frac{\phi(K_{pi})}{\sigma_p}}{\Phi(V_{pi})\frac{\phi(K_{pi})}{\sigma_p} + \left[(1-\Phi(V_{si}))\frac{\phi(K_{si})}{\sigma_s}\right]} \tag{7.7}$$

and may be calculated from the maximum likelihood estimates for each individual in the sample.

The above likelihood function is the standard stochastic switching model outlined in Maddala (1983) and used by Dickens and Lang (1985) in their application to dual labour markets. The advantages this estimation procedure possesses over the more *ad hoc* procedures that have characterised the empirical dual labour market literature are clear. An individual's sectoral attachment need not be observed by the investigator to obtain separate sectoral wage equation estimates. However, it is also clear that if one uses occupations and/or job characteristics as the criterion for allocating individuals to the primary or secondary sector some jobs can unambiguously be defined as primary and others as secondary. Though, there exists ambiguity for a large number of individuals it would be inappropriate to throw away important information from the econometric point of view if the likelihood function outlined above could be easily modified to take this information into consideration. Stewart (1988), in an application different from the one here, presents an extended likelihood function that treats the realisation of the latent dependent variable as trichotomous and partially observed. This extended likelihood function has the stochastic switching model as a special case.

In the Stewart (1988) case D defines the trichotomous realisation of the latent dependent variable y^*. If $D = 1$, the observation is in the primary sector, if 2 then the regime governing the observation is unknown and if 3 the observation falls into the secondary sector. The extended likelihood may be expressed as

$$L = \sum_{D_i=2} \log\left[\left[\Phi(V_{pi})\,\frac{\phi(K_{pi})}{\sigma_p}\right] + \left[(1-\Phi(V_{si}))\,\frac{\phi(K_{si})}{\sigma_s}\right]\right] + \tag{7.8}$$

$$\sum_{D_i=1} \log\frac{\phi(K_{pi})}{\sigma_p} + \sum_{D_i=3} \log\frac{\phi(K_{si})}{\sigma_s}$$

Thus, the log likelihood can be de-composed into the contributions made by the three different categories. The first part of the log likelihood is that used in the Dickens and Lang (1985) application and the latter two parts refer to the contributions to the log likelihood of the known primary and secondary sector's observations[66]. This likelihood is clearly much more general than the standard stochastic switching model and allows prior information to be used in the allocation of individuals to the given sectors. The prior information used to allocate individuals to the primary and secondary sector is discussed in the following section.

7.3 Data

The data used in this chapter are again obtained from a European Economic Community commissioned survey carried out by the Economic and Social Research Institute, Dublin and titled "The Transition from Education to Employment". The target group in the survey were young workers between the ages of 15 and 24 who had left full-time education and were either actively engaged in employment or actively searching for work. The sub-sample employed in this analysis is composed of those individuals of single status who defined their main economic activity as either working for payment or profit in non-agricultural activities. Only those who classified themselves as full-time workers are included. The total number of observations for which no missing values were recorded was 2831. Of these 1505 were male and 1326 female. The variables used in estimation are as follows:

Wage :Net hourly wage expressed in logarithms.

Experience: Total labour force experience expressed in years. Since the exact dates for starting and terminating jobs is known for each individual the experience variable can be calculated with a high degree of precision. For the wage equations this variable is expressed in terms of linear splines designed to capture differential returns over different parts of an individual's labour force experience. The experience variable that enters the switching equation, however, is the full variable.

---

[66] It is clear that the probabilities are defined only over the D=2 cases.

Education: Number of years in post-compulsory education. Fifteen is the school-leaving age.

Current Residence: A set of three (0,1) dummies for an individual's current region of residence. Dummies for Dublin city and county, the southern counties, the midlands and the east are included with the omitted variable in estimation being current residence in the north-western counties.

Firm Size: A set of three (0,1) dummies for the size of the firm the individual currently works in. The three dummies are for firms less than one hundred workers, firms with greater than one hundred workers but with less than five hundred workers and firms with greater than five hundred workers. In estimation the omitted dummy is firms with less than one hundred workers.

Sex. A (0,1) dummy adopting a value of 1 if male and 0 if female.

Public. A (0,1) dummy that captures whether the individual is in a public sector job or not. The public sector definition used is relatively narrow and refers mostly to those employed as civil servants. This is due to the fact that it is not possible, on the basis of the data, to allocate those working in education or health services conveniently into public or private categories[67].

For the purposes of this analysis certain types of individuals are allocated to either the primary or the secondary sectors[68] depending on their individual and/or job characteristics. The public sector variable, the firm size variables and the occupational information contained in the data set provide the prior information to be used for this allocation. These variables are used because it is felt that the more precise the dual labour market definition employed, the more precise will be the model's coefficient estimates. A strong argument must, however, be made for the use of industry information in the allocation of workers to primary and secondary type jobs since it is clearly the case that primary and secondary type jobs co-exist within a given industry. However, in the context of the public sector an exception can be made since it is clear that using this in line with occupational information a straightforward sectoral delineation can be made.

---

[67] This set of variables comprise a sub-set of variables used in the estimation. Other variables, in particular, unemployment duration and change of residence failed to record an acceptable level of statistical significance and are thus not contained in the reported estimates.

[68] As was pointed out above the primary/secondary dichotomy may be viewed as an abstract from reality that facilitates the subsequent empirical analysis.

The allocation of workers and the assumptions made to justify this allocation are as follows:

(i) All professional workers (as defined by the aggregate occupational codes 2, 3, 4 and 5 in the Census of Population Classification of Occupations (1981)) are assigned with probability one to the primary sector. This is a relatively innocuous assumption since it is clear that most professional jobs satisfy the criteria for inclusion in the primary sector, *i.e.* high returns to education, high wages, stable employment and good work conditions *etc.*

(ii) All non-manual (excluding professional workers as allocated above) public sector employees, male and female, are also allocated to the primary sector. The justification for this is that for non-manual workers, in particular, public sector employment is characterised by a distinct carrer structure with internal labour markets *etc.* and readily fits most of the stylised facts of a primary labour market. Since most manual public sector employees are in refuse disposal, cleaning services and the administration of cemetries *etc.* it would be inappropriate to allocate them to the primary sector. However, due to the existence of various fringe benefits for public sector employees it would be inappropriate to allocate them to the secondary sector. Consequently they are allocated to the unknown sector.

(iii) Only those skilled manual workers, male and female, who work in firms of greater than one hundred workers are allocated to the primary sector. The justification for this allocation lies in the fact that small firms are not viewed as being characterised by the type of internal labour markets and promotional opportunities compatible with a primary sector classification. Though it can be argued that secondary jobs co-exist with primary jobs in large organisations skilled workers are more likely to hold down the "good" jobs within these firms. Furthermore, employment stability is more likely to characterise larger rather than smaller firms. Since it would be inappropriate to allocate skilled manual workers to the secondary sector the residual is allocated to the unknown sector.

(iv) All the unskilled manual workers are allocated to the secondary sector. A number of semi-skilled, intermediate non-manual and other non-manual jobs are also allocated to this sector. The semi-skilled occupational codes allocated to the primary sector are 290, 291, 292, 293, 313, 340 and 342. The intermediate non-manual occupational codes allocated to the secondary sector are 325 and 326 and the other non-manual codes are 330, 331, 336, 337 and 339[69]. The semi-skilled workers allocated to the secondary sector include dock labourers, drivers' helpers and charwomen and cleaners *etc.* with the other semi-skilled and non-manual workers allocated to the unknown category[70]. Since an individual's sectoral attachment is on the basis of their characteristics and the characteristics of their job all the variables that enter the wage equations also enter the sectoral attachment equation. However, the experience variable enters the switching equation as a continuous variable rather than as two piece-wise linear segments as is the case in the wage equations. Furthermore, in the case of the two wage equations the sex dummy variable is allowed to interact with the experience splines. Other interactive effects were also used in estimation but with little effect[71].

It should also be pointed out at this stage that the estimates reported in the subsequent subsection appear relatively robust to changes in the prior allocation of certain workers. Experimentation involved reducing and increasing the firm size threshold for the allocation of skilled workers to the primary sector. Other experimentation involved the allocation of all public sector employees as defined above to the unknown sector, the allocation of all the semi-skilled workers to, in turn, either the unknown or the secondary sector. The allocation of non-manual female public sector employees to the unknown sector. The allocation of all skilled workers regardless of firm size to the primary sector. The results obtained with the actual allocation reported above proved relatively robust to the above modifications.

7.4 Empirical Results

Table 7.1 reports the empirical estimation using maximum likelihood methods of equations (7.1) to (7.3). For purposes of comparison with the wage equation estimates OLS estimates are

---

[69] These non-manual categories include shop assistants, roundsmen, waiters and waitresses, domestic servants, street vendors. These codes are contained in the Census of Population Classification of Occupations (1981) cited above and a more extensive treatment appears in appendix 7.A1 of this chapter.

[70] The final numbers allocated to the sectors are 694 to the primary sector, 619 to the secondary sector and 1518 to the unknown sector.

[71] The other variables in the specification that sex was interacted with but to little statistical effect were Dublin, public sector employment and education.

also reported. The OLS estimates refer to the estimates from a full sample equation that ignores the partition. Attention initially focuses on the switching coefficients. As outlined in section 7.2 it is not possible to identify all the coefficients of the switching equation and its error variance simultaneously. Thus a normalisation of unity is used. This has obvious implications for the interpretation of the switching coefficient estimates. Since the signs of these coefficients are not affected by the normalisation they are useful in providing information on sectoral attachment. All the variables' coefficients are negative in sign. If the results are taken to suggest secondary sectoral attachment they are clearly consistent with one's priors regarding dual labour markets. The switching coefficients could be interpreted in summary to suggest that the more education one possesses, the larger the firm size one works in, being resident in Dublin, being in the public sector and being male all reduce the probability of secondary sector attachment. These findings could be interpreted as supportive of the substantive predictions of the SDLM hypotheses. Furthermore, some measure of comfort is provided by the lack of conflict in these results.

A distinguishing feature between the estimates reported for the two sectors is the private rate of return to education. As SDLM theorists suggest this should be higher in the primary relative to the secondary sector. This is indeed the case where the returns to education are 8.7% in the primary sector and only 5.1% in the secondary sector. The implications of this set of estimates for the wage differentials are discussed more extensively in section five[72].

Though it is not possible to compare either the firm size or the regional coefficients across the two sectors it is interesting to note that secondary workers in terms of both the firm and regional variables are better off than primary workers relative to their respective base groups. Both these sets of results could be interpreted as suggesting a wider variance in net hourly wages in secondary than in primary jobs. This could also be interpreted as suggesting that individuals in the secondary sector who work in the smallest sized firm and reside in the north/western region are much worse off in relation to other secondary sector workers than are primary workers in these particular categories. The public sector coefficient in statistically significant in the primary

[72] Suffice for it to state here that there appears little evidence to suggest that the wage profiles in the secondary sector are flat in relation to years in post-compulsory education.

sector suggesting that individuals in primary public sector employment receive a wage that is 11.0% higher than comparable individuals in the primary private sector earn.

Attention now turns to the interpretation of the experience terms. The linear splines allow for differential returns over two parts of a young workers' labour force experience. The interaction of sex with the linear splines allows for these returns to be decomposed by gender. The inclusion of the interaction terms alters the interpretation of the spline coefficients. These now represent the returns to females of the different years labour force experience and the coefficients on the interaction terms are the male/female differentials in these returns. The coefficient on the sex dummy now reflects the returns to being male at zero experience, *ceteris paribus*. It is clear on inspecting the results that the returns to female labour force experience in the first four years is only marginally higher in the primary sector and on the basis of a t-statistic are not statistically different from the returns that obtain in the secondary sector. The same can be said, on the basis of a similar t-test, for returns in those years subsequent to the fourth. In the primary sector the gender differential in the first spline of experience is statistically significant with the second spline less so. In contrast, neither of the interactive terms in the secondary sector's wage equation are statistically significant. Again on the basis of t-statistics the gender differential in the first spline of experience in the primary sector is not found to be statistically different from that found in the secondary sector. No statistical effect is detected for the spline associated with the latter years of labour force experience[73].

Interestingly enough, however, at zero labour force experience females employed in the primary sector earn approximately 17% more than comparable males, *ceteris paribus*. This clearly conveys something about the types of jobs males and females do in the primary sector and a rationalisation for this result may be found in chapters five and six. A large number of the males allocated to the primary sector with probability one belong to the skilled manual sector. Since they undergo firm-specific training their starting wages are considerably lower than that for the average in the sample (see chapter six). Since females tend to do jobs in the primary sector that

---

[73] The t-statistics (in absolute values) for the tests of primary/secondary differences in labour force experience are the order of sequence cited in the text 1.057, 0.3831, 1.533 and 0.756.

do not involve such short-term sacrifice their starting wages are likely to be higher. This result should not be interpreted as being inconsistent with the predictions of the dual labour market but as being more consistent with the fact that within the primary sector certain characteristics of male jobs are different from those of females. In contrast starting wages for males in the secondary sector are higher by nearly 7%[74].

It may be more instructive to graphically illustrate the wage/experience profiles for both males and females in the primary and secondary sectors respectively. Figure 1 in appendix 7.A2 of this chapter displays the male and female wage/experience profiles for the primary sector coefficient estimates. It obviously confirms most of the interpretations cited above. For both sexes the returns to labour force experience are steeper in the early years of labour force experience. This fact has been noted in all of the previous empirical chapters. It is also clear from figure 1 that in the early years of experience the female wage profile is higher than the corresponding male profile. Job specific training on the part of males may be offered as a rationalisation for this occurrence[75]. The cross-over point is somewhere near the third year of labour force experience. Subsequent to this the male profile always exceeds the female profile and the perceptible gap between the two profiles is seen to widen with the addition of labour force experience. This could well be explained by the fact that males are more likely to gain promotion and ascend much more rapidly than females the internal ladders assumed to characterise the primary sector. In fact after the fourth year of labour force experience the female profile has a relatively flat configuration. In summary figure 1 is not that much in conflict with the predictions of human capital theory.

Figure 2 tells a slightly different story for the secondary sector. In this case male wages are always higher than female wages and the gap in profiles is, as in the primary sector, seen to widen with the passage of labour force experience. As was pointed out above it is difficult to provide a

---

[74] This could be indicative of discrimination effects operating in in the secondary labour market. However, there is a view (see Wachter (1974)) that secondary jobs are concentrated in those firms that possess in Becker's terminology low discrimination coefficients. If secondary sector wages are determined by competitive forces little latitude exists for employers to discriminate. Thus this result appears to mildly conflict with some theoretical predictions.

[75] See chapter six.

theoretical explanation for this result couched exclusively in discriminatory terms. The predictions of dualism that suggest a relatively flat wage profile in regard to experience are evidently not verified in the data provided by this particular sample of young workers.

On the basis of the coefficient estimates and Bayes' theorem outlined in (7.7) it is possible to allocate the 1518 unknown individuals to the primary and secondary sectors on the basis of their predicted conditional probabilities. Since the switching equation coefficient estimates are in line with predicting the probability of secondary sector attachment a conditional probability of greater than 0.5 allocates an individual to the secondary sector. If the predicted conditional probability is less than 0.5 the individual is allocated to the primary sector. On the basis of this rule of thumb 49% from the unknown category are allocated to the secondary sector with the remainder allocated to the primary sector. Thus as Table 7.2 indicates the total in the primary sector is now 1464 with 1367 in the secondary sector. Table 7.2 also contains the sample means for the variables used in estimation for the full sample and those allocated to the primary and secondary categories. An inspection of table 7.2 reveals an observed wage differential between the two sectors of over 30%. The sectoral gender wage gaps (featured in table 7.3) are 3.7% and 2.0% for the primary and secondary sectors respectively. This compares with a gender wage differential of 3.0% for the sample as a whole. Labour force experience and education are on average higher in the primary sector and only 16% of secondary workers work in firms with more than one-hundred workers. The comparable primary sector figure is nearly 70%. Significantly more primary workers reside in Dublin and over 20% of this sector's workers are in the narrowly defined public sector. The sectoral allocation by gender is, however, reasonably comparable.

Table 7.4 provides means for the primary and secondary sector broken down by gender. These means are not substantially different from those reported and commented on in table 7.2. However, one interesting feature of table 7.3 is the fact that nearly 27% of all females in the primary sector are in public sector employment.

The single equation model reported in column 1 of table 7.1 is nested within the switching model. It is obtained by the imposition of the restrictions, $\beta_p = \beta_s$, $\sigma_p = \sigma_s$ and $\rho_p = \rho_s$. A likelihood ratio test may thus be used to statistically test whether these restriction are upheld by the data. Since several parameters are unidentified the calculation of the appropriate degrees of freedom is difficult. Goldfeld and Quandt (1976) suggest setting the number of degrees of freedom equal to the number of restrictions to be tested plus the number of unidentified parameters. In terms of table 7.1 this generates 26 degrees of freedom for the χ^2 of the likelihood ratio test. The calculated test statistic is 273.6 and with a critical value of 45.6 at the 1% level of significance the null hypothesis of the single equation model is decisively rejected by the data[76].

---

[76] The null hypothesis in this case is that a labour market is characterised by a continuum of jobs. This is consistent with the competitive human capital model. The alternative hypothesis is a labour market segmented into primary and secondary sectors on the basis of the taxonomy outlined. It is clear that the degrees of freedom used provide a very stiff test for the hypothesis under consideration. Though this test has been used by Dickens and Lang (1985) it should be pointed out that under the null the ρ's are not estimated.

Table 7.1

Endogenous Switching Coefficient Estimates.

Variable	OLS	Primary	Secondary	Switch
Constant	-0.1773***	-0.2340***	-0.2218***	1.3418***
	(0.0281)	(0.0455)	(0.0365)	(0.3445)
Education	0.0695***	0.0869***	0.0509***	-0.1358**
	(0.0046)	(0.0068)	(0.0066)	(0.0608)
100 ≤ Firm < 500	0.1412***	0.1528***	0.1533***	-0.9137***
	(0.0141)	(0.0294)	(0.0258)	(0.2366)
Firm ≥500	0.1672***	0.1823***	0.1915***	-0.5373
	(0.0146)	(0.0265)	(0.0248)	(0.2441)
Resident in Dublin	0.1512***	0.1304***	0.1706***	-0.4099***
	(0.0163)	(0.0313)	(0.0274)	(0.0269)
Resident in South	0.0726***	0.0725**	0.0746***	-0.0449
	(0.0181)	(0.0307)	(0.0266)	(0.2783)
Resident in East & Mid	0.0385**	0.0396	0.0387*	-0.0150
	(0.0187)	(0.0327)	(0.0275)	(0.2590)
Public	0.1038***	0.1047***	0.0760	-1.4315*
	(0.0166)	(0.0292)	(0.0468)	(0.7722)
Sex	-0.0844**	-0.1551***	0.0669*	-0.0390
	(0.0336)	(0.0462)	(0.0377)	(0.1807)
Exper	-	-	-	-0.2184***
	-	-	-	(0.0495)
Exper. ≤ 4 years	0.0837***	0.0953***	0.0769***	-
	(0.0080)	(0.0129)	(0.0099)	-
Exper. > 4 years	0.0310***	0.0313	0.0409***	-
	(0.0083)	(0.0214)	(0.0096)	-
Sex ×Exper≤4	0.0464***	0.0508***	0.0150	-
	(0.0114)	(0.0172)	(0.0141)	-
Sex ×Exper>4	0.0243**	0.0397*	0.0172	-
	(0.0106)	(0.0253)	(0.0128)	-
σ	0.3003	0.3309	0.2663	**#**
	-	(0.0071)	(0.0062)	-
L	-604.93		-468.12	

Asymptotic standard errors are in parentheses and * denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. The number of observations is 2831 and bold # denotes normalisation to unity. σ is the standard error of the equation.**

Table 7.2

Means of Variables and Sectoral Allocation of Workers

Variable	Full Sample	Primary	Secondary
Known	-	694	619
Unknown	-	770	748
Total	2831	1464	1367
Means of Variables			
Wage	0.3993	0.5499	0.2380
Experience	3.326	3.942	2.667
Education	1.902	2.222	1.559
100 ≤ Firm < 500	0.1692	0.2725	0.0585
Firm ≥500	0.2741	0.4296	0.1075
Current Residence			
Dublin City & Co.	0.3497	0.4590	0.2326
South	0.2882	0.2541	0.3248
East & Mid.	0.1982	0.1667	0.2319
Sex	0.5320	0.5307	0.5333
Public	0.1194	0.2077	0.0249
Exper. ≤ 4 years	2.756	3.126	2.360
Exper. > 4 years	0.5704	0.8167	0.3066
Sex ×Exper≤4	1.499	1.675	1.310
Sex ×Exper>4	0.3316	0.4837	0.1687

The unknown workers are allocated to the secondary sector on the basis of whether the conditional probability (see equation (7)) > 0.5. The others are allocated to the primary sector.

Table 7.3

Sectoral Means by Gender

Variable	Primary Male	Primary Female	Secondary Male	Secondary Female
Means of Variables				
Wage	0.5672	0.5304	0.2474	0.2273
Experience	4.067	3.802	2.774	2.544
Education	2.071	2.393	1.362	1.784
100 ≤ Firm < 500	0.2188	0.3333	0.0576	0.596
Firm ≥500	0.4414	0.4163	0.0919	0.1254
Current Residence				
Dublin City & Co.	0.4363	0.4847	0.2071	0.2618
South	0.2600	0.2475	0.3100	0.3417
East & Mid.	0.1828	0.1485	0.2483	0.2132
Skilled	0.4492	0.0684	0.4266	0.0407
Public	0.1532	0.2693	0.0302	0.0188
Number	777	687	729	638

The unknown workers are allocated to the secondary sector on the basis of whether the conditional probability (see equation (7)) > 0.5. The others are allocated to the primary sector.

7.5 Dual Labour Market Wage Differentials

Table 7.1 highlighted some differences across the two sectors in terms of returns to key variables. This section attempts to examine this issue in more depth by the examination of mean wage differentials. The differentials are calculated on the basis of the means of the full sample of individuals. This seems an appropriate approach to adopt since it purports to ask the primary/secondary differential in wages that would confront an individual with a mean set of characteristics. On the basis of equations (7.1) and (7.2) this differential may be expressed as

$$\hat{\bar{\psi}} = \bar{X}\hat{\beta}_p - \bar{X}\hat{\beta}_s \tag{7.9}$$

where $\hat{\bar{\psi}}$ is the mean differential estimate with p and s denoting primary and secondary respectively, carats denoting estimates and bars means. More conveniently (7.9) may be expressed as

$$\hat{\bar{\psi}} = \bar{X}\Delta\hat{\beta} \tag{7.10}$$

The variance for this expression may be written as:

$$\mathrm{var}(\hat{\bar{\psi}}) = \bar{X}'V\bar{X} \tag{7.11}$$

where

$$V = \mathrm{var}\,(\hat{\beta}_p) + \mathrm{var}\,(\hat{\beta}_s) \;-\; 2\,\mathrm{cov}\,(\hat{\beta}_p\,\hat{\beta}_s)$$

The asymptotic standard error is thus given by $\sqrt{\mathrm{var}(\hat{\bar{\psi}})}$ and the consequent test statistic is asymptotically normally distributed under certain assumptions *e.g.* normally distributed errors.

The mean differential calculated as *per* (7.10) is reported in table 7.4 and suggests a differential in favour of the primary sector of 3.6% which is also found to be statistically significant at the 10% level using a two-tailed test. This implies that given a mean set of characteristics their valuation is almost 4% higher in the primary relative to the secondary sector and suggests the existence of two distinct wage setting mechanisms that pay a given set of characteristics differently. A differential for a stylised individual with mean labour force experience, mean education and with a base set of the other characteristics[77] is also calculated and is reported as 12.5% and again is statistically significant.

In order to assess how the differential varies across different categories differentials are also reported for the deviations from this base set of characteristics. Firm size, being resident in the southern or east and midlands regions and being in the public sector all record statistically significant effects which in most cases are well over 10%. However, surprisingly, being resident in Dublin and being male possesses no statistical advantages in terms of the dual labour market differentials.

The most crucial variables to the SDLM story are the education and experience variables. Differentials by sector based on deviations from the mean by education and experience are also presented in table 7.4. The stark feature of this table, but not surprising in view of the coefficient estimates of table 7.1, is the effect of education on the sectoral wage gap. For example, with six years post-compulsory education the differential reported is almost 20% confirming the importance of this variable in the determination of a sectoral wage gap. The wage gap is seen to widen dramatically as the years in post-compulsory education increase. Incrementing the experience variable also has a widening effect on the differential, albeit less dramatic. Nevertheless, both variables confirm a basic tenet of the dual labour market hypothesis *i.e* identical levels of education and experience are remunerated at a significantly higher rate in the primary relative to the secondary sector. However, having said this it is also clear from the estimates of table 7.1 that there is no evidence of a relatively flat wage profile in the secondary sector in terms of either experience or education and thus a more important tenet of the dual labour market hypothesis is not confirmed by the data.

Finally, the last two rows of 7.4 examine the primary/secondary sector wage differential for a stylised male and female individual. For a male with five years work experience, five years post-compulsory education, working in the largest firm size, residing in Dublin and working in

[77] The stylised individual with mean education and experience in this case works in the smallest sized firm, resides in the north/western region, works in the private sector and is female.

the public sector the calculated differential is 17.0%. The estimate for a comparable female is 23.5%. On the basis of this particular exercise it is clear that females in the secondary sector are at a clear disadvantage. As pointed out in discussing the results this may be explained by their relatively low pay status in the secondary sector.

Table 7.4

Primary/Secondary Wage Differentials by Category.

Variable	$\bar{X}\Delta\beta$	ase
Mean	0.0353*	0.0178
Base	0.1175**	0.0490
Deviations from Base		
100 ≤ Firm < 500	0.1171**	0.0576
Firm ≥500	0.1084**	0.0576
Dublin City & Co.	0.0773	0.0503
South	0.1154**	0.0426
East & Mid	0.1185**	0.0461
Sex	0.0184	0.0450
Public	0.1463**	0.0695
Deviations from Mean by Education		
Education = 1 year	0.0004	0.0199
Education = 2 years	0.0365**	0.0178
Education = 3 years	0.0725***	0.0212
Education = 6 years	0.1807***	0.0460
Deviations from Mean by Experience		
Experience = 5 years	0.0787***	0.0246
Experience = 6 years	0.0811***	0.0304
Experience = 7 years	0.0835**	0.0415
Male#	0.1587***	0.0634
Female#	0.2114***	0.0688

ase denotes asymptotic error, *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level using two tailed tests. $\Delta\beta$ denotes difference between the primary sector and secondary sector wage coefficient estimates. Deviations from the base or mean occur singly. # denotes five years work experience, five years of post compulsory education, working in the largest firm size, residing in Dublin and working in the public sector.

7.6 An Informal Test of Rationing

As pointed out in the introductory section the formal test of rationing proposed by Dickens and Lang (1985) has been criticised as uninformative and dismissed on econometric grounds by Heckman and Hotz (1986). In any event their test would be inappropriate in terms of the framework employed here. Thus as a very informal test of rationing sectoral differentials are calculated for all individuals in the sample on the basis of (7.10) with appropriate standard errors[78]. This allows those individuals allocated to the secondary sector (either by assumption or by the switching model) who have a positive and statistically significant ψ to be identified. The existence of a positive differential for secondary workers may be viewed, if one assumes supply side neutrality, as reflecting rationing behaviour on the part of employers[79]. It is also clear that a similar exercise can be effected for the primary sector's workers.

Table 7.5 reports estimates based on such an exercise with a critical asymptotic t-value of 1.96 used as the cut-off point between statistical significance and insignificance. As column two of table 7.5 reveals 140 of the 1367 individuals allocated to the secondary sector had a statistically significant ψ. These results suggest that over 10% of those in the secondary sector would earn more (on average 14.0% more) in the primary sector than what they actually earn in the secondary sector. An examination of the characteristics of this particular sub-set of workers shows that they possess a mean education well above the average for the sample and that nearly all are female. A tentative dual labour market interpretation for this result is that employers use gender to restrict female entry into certain primary jobs. The result, however, is in no way inconsistent with a supply side interpretation that is explained in terms of female preferences for the non-pecuniary benefits of secondary sectoral attachment. In other words females may prefer the flexible working arrangements, the low-level responsibilities and the temporary nature of secondary jobs because of their anticipated marginal adherence to the labour market. A similar exercise is carried out for the primary sector workers revealing that nearly 14% of these possess

[78] This informal test is an extension of Stewart (1983).

[79] Though this assumption may be viewed as a relatively strong one it is no stronger than the assumptions made concerning the empirical counterpart of the Becker discrimination coefficient in chapters four, five and six of this thesis.

differentials that are statistically significant. The results of this informal test appear to suggest that the primary/secondary dichotomy only matters for less than 25% of the total sample. For the remainder in the sample there exists a statistical indifference as to their sectoral allocation. That is to say over 75% of the sample would be neither worse nor better off in primary as opposed to secondary (or vice-versa) job status.

Table 7.5

Sectoral Differentials by Individual

Variable	$\psi_i > 0^a$	$\psi_i > 0^b$
Means of Variables		
Wage	0.5310	0.3714
Experience	4.112	2.121
Education	3.347	3.093
$100 \leq$ Firm < 500	0.2764	0.0071
Firm ≥ 500	0.3317	0.0357
Current Residence		
Dublin City & Co.	0.3166	0.0786
South	0.3317	0.4143
East & Mid.	0.1960	0.2857
Sex	0.1206	0.0143
Public	0.0000	0.0000
$\bar{\psi}$	0.1455	0.1316
Number	199	140

Standard errors are calculated as in the text. The significance level chosen is 5% and using a two-tailed test the |t| value is 1.96. The primary sector is denoted by a and the secondary sector by b. Switch denotes the numbers allocated to the category by the switching model

7.7 Conclusions

This chapter has attempted to establish the existence of two separate wage setting mechanisms that are characterised by the predictions of dualism. An endogenous switching econometric model with a trichotomous realisation of the latent dependent variable (and which possesses the stochastic switching model as a special case) was used in estimation. The single equation model was decisively rejected against the two sector alternative. Some of the results obtained could be rationalised in terms of what dualism predicts. A dramatic difference in the private rate of return to education was observed across the two sectors. Education in particular was seen to play a strong role in the determination and widening of the wage gap between the two sectors with labour force experience playing a less significant role. Nevertheless, no evidence of a flat wage profile with respect to either experience or education was detected for the secondary sector.

Tentative evidence of a male wage advantage in terms of the primary and more strongly in terms of the secondary sector was detected. However, for those females in the unknown sector gender appeared to play an insignificant role in terms of their sectoral allocation. The gender distribution across sectors was in line with the sample mean and there was no evidence that females were over-represented in the secondary sector. The youth and marital status of the sample may be invoked to provide a rationalisation for this particular result. For the UK Dex and Shaw (1986) and Stewart and Greenhalgh (1984) provide evidence of downward female occupational mobility after labour force interruption. It could thus be suggested that the segregation of females into secondary type jobs would not be detected in a sample of females who have experienced no such discontinuity in participation. Nevertheless, on the basis of an informal test of rationing the most disadvantaged group in the secondary sector was females. Nearly, all of those in the secondary sector who could earn significantly more in the primary sector were female.

From the perspective of the sample of individuals as a whole only slightly under 25% had a statistically significant positive ψ. Thus, almost three-quarters of the sample are indifferent to either a primary or a secondary sector allocation. The youth of the sample can again be invoked to provide an explanation for this particular result. It is well known that young workers in trying to establish their optimal job match sample a number of jobs prior to settling down. They thus may straddle the primary/secondary sector divide. As highlighted in 7.4 the widest differentials only appear for those young workers who are near the terminal age for those in the sample. It is conceivable that some young workers are optimising in terms of the short-run and thus take jobs in the secondary sector that have higher starting wages in preference to a primary job that has better long-run prospects.

Since 75% of individuals are indifferent as to their sectoral allocation the effectiveness of the switching equation must also be brought into question. It could, however, be said that the task set the switching model is far too great given the youth of the sample under consideration. If there is little evidence of dualism in the young segment of the labour market, the switching equation will have severe problems in successfully delineating the two sectors. The more appropriate question is whether the application of this theoretical model to the young workers' labour market is valid. On the face of the results presented in this chapter little evidence of dualism can be discerned.

Appendix 7.A1

Occupation and Industry Codes of Allocated Workers

The Census of Population Classification of Occupations (Central Statistics Office, Dublin (1981)) is used to allocated the workers with probability one to either the primary or the secondary sectors.

The following occupational codes are allocated to the primary sector: 329, 355, 357, 358, 359, 360, 365, 366, 367, 369, 370, 373, 377, 378, 380, 381, 382, 383, 396, 384, 395, 356, 361, 368, 371, 372, 374, 375, 376, 379, 385, 386, 388, 389, 390, 392, 393, 318, 319, 320, 321, 322, 323, 333, 334, 344, 351, 352, 354, 394, 309, 324, 327, 328, 332, 363, 364, 397.

The following occupational codes are allocated to the secondary sector: 290, 291, 292, 293, 313, 340, 342, 294, 295, 296, 325, 326, 330, 331, 336, 337, 339.

Furthermore, all non-manual public sector employees in industry codes (SIC) 911, 912, 913, 914, 915, 919, 921, 922, 923, 931, 932, 933, 000, 951, 952 are allocated to the primary sector.

In addition all those workers in the one digit skilled manual occupational category and who work in firms with over one hundred workers are also allocated to the primary sector.

All other occupational and industry codes not already classified are allocated to the unknown sector.

Figure 1.

Primary Sector.

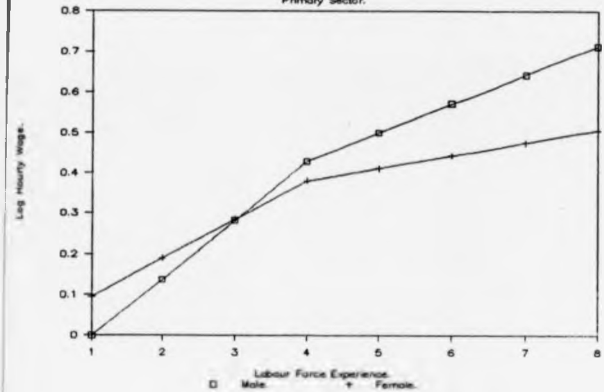
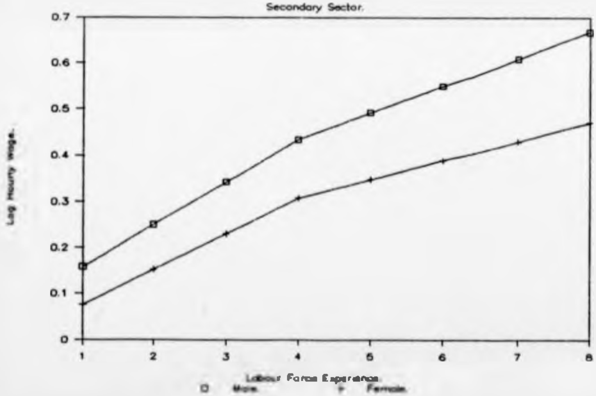


Figure 2.

Secondary Sector.



Appendix 7.A2

Chapter Eight

Conclusions

It is a difficult exercise to draw the diverse strands of the previous four empirical chapters together and present an overall set of conclusions. The exercise may, therefore, best be served by examining in turn the findings of these chapters while identifying the linking themes.

Firstly, chapter four presented, to the author's knowledge, the first set of wage equation estimates for young workers in Ireland. The equations estimated provided an opportunity to assess the wage effects of labour force experience, formal educational and vocational training qualifications. The results also provided an insight into the influence exerted on wages by certain job characteristics, occupation, industry and regional location. Most of the results reported were found to be consistent with one's priors.

A more important objective of this particular chapter was the estimation of gender discrimination coefficients using the "index number" approach. Mean discrimination estimates ranged from a statistically insignificant 3% to nearly 9% depending on whether the equations in question conditioned on occupations and industries or not. Thus the larger estimates reported may say more about the effects of occupation and industry compensating differentials than about discrimination. The evidence presented in this chapter provided little indication of the existence of a gender based wage discrimination.

Chapter five provided an advance on chapter four in terms of its treatment of occupations. The previous chapter assumed occupational exogeneity and the estimates presented were based on this rather strong assumption. The main purpose of chapter five was an analysis of the effects of occupational endogeneity on the wage equation estimates and hence the discrimination estimates. In this chapter occupations were dichotomised on the basis of non-manual and manual workers. Two contrasting econometric methods were employed to estimate the occupational wage equations. One was the standard Heckman two-step estimator and the other was a more recently developed distribution free instrumental variable (IV) estimator. Statistical tests for occupational exogeneity were provided using both these estimators and some conflicting results were reported. The IV based Hausman test revealed no evidence of occupational endogeneity in either male or female equations. In contrast the Heckman based Melino test provided some evidence of occupational endogeneity in the manual male equation.

Not surprisingly the gender wage discrimination effects proved sensitive to the estimator used. Both estimators revealed little evidence of manual wage discrimination. This may be due to the large number of skilled male workers who in the early years of labour force experience are assumed to sacrifice wages to undergo firm-specific training. However, the discrimination coefficients for the female intensive non-manual category suggested the existence of wage discrimination. The disagreement between the estimators occurred in terms of the reported magnitude which ranged between just over 8% to under 16%.

Finally, in this chapter estimates for a structural occupational model were also reported and suggested, in general, an unimportant role for the occupational wage gap variable in terms of non-manual occupational attachment. Of much more importance for both gender groups were parental background, education, residence of schooling and to a lesser extent the duration of unemployment since leaving school.

A limitation in chapter five was the inability created by the very broad occupational dichotomy to assess the effects of occupational segregation on the gender wage gap. Chapter six examined an extension to the "index number" approach that allowed the gender wage gap to be decomposed into intra and inter occupational components. A five-way occupational categorisation was employed and the information obtained from estimating reduced form multinomial logit occupational models was used to correct the occupational wage equations for potential selectivity bias. Lee (1983) suggests an econometric methodology that provides consistent estimates in the polychotomous framework. These estimates were contrasted with standard OLS ones. Evidence of the non-random allocation of workers to certain occupations was detected for a number of categories for both gender groups. In particular a strong negative selectivity bias effect was detected for both males and females in terms of the skilled occupational category re-enforcing some of the results obtained in chapter five by the Heckman based Melino test.

The methodology employed allowed for the calculation of "cleaner" occupational wage discrimination effects based on a much finer occupational classification than those reported for chapter five. Discrimination coefficients for the OLS and the consistent estimates were calculated. As was made clear in both chapters five and six though the interpretation of the OLS and the consistent differentials are different some large intra-occupational differentials were detected. This served to illustrate how the mean estimate can obscure the existence of relatively large differentials within certain occupations.

Chapter six also simulated female occupational distributions assuming a male occupational structure. In general it found that females were over-represented in the low-skill, low-pay sectors and under-represented in the skilled and professional sectors. The major finding of chapter six, however, was that the occupational allocation of females was not the major determinant of their wage disadvantage. In line with evidence presented by Brown *et al.* (1980) and Miller (1987) for the US and the UK respectively intra-occupational wage disadvantage was a much stronger determinant of the gender wage gap than the occupational allocation of females. The amelioration of the female position, it was concluded, could best be served by promoting equal pay within occupations rather than designing policies aimed at gender equalisation across occupations.

The final empirical chapter, chapter seven, attempted to use *inter alia* occupational information to help describe two sectors that embodied some of the characteristics of the segmented or dual labour market theory. An endogenous switching model with a trichotomous realisation of the latent dependent variable was used to estimate the parameters of the dual labour market model. Some of the results could be rationalised in terms of the predictions of dualism with education in particular playing a major role in both the determination and widening of the sectoral wage gap. Though the gender wage/experience profiles in the primary sector conformed to theoretical predictions the comparable profiles in the secondary sector exhibited increasing returns. A discernible gender gap between the two profiles was also observed. Since secondary firms are assumed to be characterised by competitive conditions the existence of a gender wage gap in the secondary sector was, to say the least, surprising.

The role of gender in the dual labour market was more extensively examined in the context of an informal test of rationing. Rationing of primary sector jobs is a basic tenet of dualism and the results reported suggested that most of those individuals in the secondary sector who could earn more in the primary sector were female. Though it is arguably the case that females may possess a greater preference for the low-skilled and low paid jobs in the secondary market the preferences of employers and their rationing behaviour cannot be over-looked. Nevertheless, a strong result of chapter seven was the fact that for over 75% of the sample an indifference existed between primary and secondary type jobs. This forces the conclusion that in a sample of young workers with limited work experience it may be very difficult for all but a small sub-sample to delineate two distinct sectors.

One overall conclusion that emerges from an examination of chapters four, five and six is the sensitivity of the discrimination estimate to occupational disaggregation. These three chapters could be viewed as an exercise in the progressive disaggregation of the occupational variable. It is clear from a simple inspection of the OLS results that the highly aggregative treatment of occupations can lead to misleading interpretations regarding the magnitude of the discrimination effects. The negligible discrimination estimates of chapter four masked a relatively large and significant unexplained differential in the non-manual sector, as detected in chapter five. This, in turn, masked an even larger unexplained differential, as observed in chapter six, for the non-clerical non-manual sub-set of workers. Thus, if one assumed away the potential problems associated with using OLS in these particular instances, these results could be interpreted as providing worthwhile evidence as to the dangers associated with the use of highly aggregated occupational data.

However, the problems of occupational endogeneity cannot be conveniently assumed away, as the results of chapters five and six confirm. The evidence of both these chapters indicated the sensitivity of the discrimination estimates to the manner in which occupational endogeneity is controlled. The relative merits of the IV and Heckman-type two-step procedures were explored in chapters five and six. Though the latter procedure provided an alternative interpretation for the unexplained differential, in terms of unexplained wage offer differentials, decisions regarding the econometric advantages of one estimator over another remain unresolved. Finally, the evidence of chapter seven appears to suggest that the application of a dual labour market model, in the particular context of young workers, may be an inappropriate exercise.

Appendix A1.

Table A1.1.

Mean Values of Variables for Chapter Four

Variable	Male	Female
ln (Wage)	0.5866	0.5089
Experience in years	4.2820	3.2486
Educational Qualifications		
Intermediate Certificate	0.3140	0.2112
Leaving Certificate	0.2683	0.6562
University Degree	0.0223	0.0087
Other Qualifications		
Group Certificate	0.1938	0.0192
Apprenticeship	0.3720	0.0558
Basic Training Qual.	0.4165	0.6771
Job Characteristics		
Estab < 15	0.3318	0.3192
15 ≤ Estab < 50	0.1626	0.1571
50 ≤ Estab < 100	0.0980	0.0890
Estab ≥ 100	0.4076	0.5096
Promotion on Job	0.4632	0.4014
Occupations		
Professional	0.0512	0.0593
Self-employed	0.0312	0.0174
Salaried employees	0.0200	0.0052
Intermediate non-manual	0.1559	0.7103
Other non-manual	0.0757	0.0942
Skilled manual	0.5501	0.0349
Semi-skilled manual	0.0779	0.0750
Unskilled	0.0380	0.0037
Industries		
Building & Engineering	0.1559	0.0140
Transport & Communication	0.0735	0.0454
Banking & Insurance	0.0557	0.2024
Public Admin. etc.	0.0891	0.2967
Metal Manufacturing	0.1336	0.0820
Other Manufacturing	0.1693	0.1344
Ext. & Chemicals	0.0668	0.0314
Distributive Trades	0.2561	0.1937
Regions		
Dublin County	0.3541	0.4537
North-West counties	0.1462	0.1153
Southern Counties	0.2695	0.2792
Midlands Counties	0.1470	0.0977
Leinster(excl. Dublin)	0.0846	0.0541
Urban	0.6147	0.6911
Observations	573	449

In terms of dummy variables the values reported represent the sample proportions in the relevant category.

Table A1.2

Means for Males and Females by Occupational Category in Chapter Five.

Variable	Male[a]	Male[b]	Female[a]	Female[b]
ln (Wage)	0.4853	0.3685	0.3949	0.3517
	(0.3722)	(0.4012)	(0.3783)	(0.2773)
Previous Exp. (years)	0.9854	0.9869	0.7834	0.9796
	(1.584)	(1.598)	(1.936)	(1.616)
Experience (years)	3.322	3.517	3.105	3.482
	(2.019)	(2.127)	(1.936)	(2.258)
Education (years)	2.310	1.376	2.407	1.139
	(1.696)	(1.288)	(1.387)	(1.215)
Schooling in Dublin	0.4577	0.2327	0.3717	0.2685
	(-)	(-)	(-)	(-)
50 ≤ Firm <400	0.2025	0.2412	0.1974	0.5586
	(-)	(-)	(-)	(-)
Firm ≥ 400	0.3556	0.2220	0.3026	0.1944
	(-)	(-)	(-)	(-)
Unemployment (in months)	2.838	4.009	2.173	3.887
	(4.466)	(7.268)	(4.065)	(6.521)
Move Residence	0.0792	0.0427	0.0922	0.0278
	(-)	(-)	(-)	(-)
Father's Occupation				
Non-Manual	0.3627	0.4739	0.3687	0.5772
	(-)	(-)	(-)	(-)
Manual	0.4842	0.2391	0.3697	0.2346
	(-)	(-)	(-)	(-)
λ	0.8560	-0.5213	0.2935	-0.9041
	(0.3797)	(0.2580)	(0.2774)	(0.5281)
Observations	568	937	998	324

a denotes the non-manual sector and *b* the manual sector. Standard deviations for the continuous variables are in parentheses.

Table A1.3.

Male Means by Occupational Category in Chapter Six.

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof & Man.
ln (Wage)	0.3724	0.3667	0.3986	0.4810	0.5961
	(0.4046)	(0.4000)	(0.3658)	(0.3497)	(0.4047)
Previous Experience	1.210	0.8932	1.335	0.7786	1.066
	(1.710)	(1.540)	(1.869)	(1.330)	(1.707)
Experience	3.452	3.544	3.728	3.197	3.144
	(2.223)	(2.087)	(2.186)	(1.928)	(1.979)
Education	1.260	1.424	1.452	2.236	3.476
	(1.273)	(1.292)	(1.375)	(1.384)	(2.023)
Residence of Schooling					
Dublin City & Co.	0.2058	0.2439	0.4178	0.4797	0.4524
	(-)	(-)	(-)	(-)	(-)
East & Midlands	0.2744	0.2258	0.3630	0.1622	0.1667
	(-)	(-)	(-)	(-)	(-)
South	0.2960	0.3273	0.2260	0.2297	0.2619
	(-)	(-)	(-)	(-)	(-)
Current Residence					
Dublin City & Co.	0.2022	0.2439	0.4384	0.5101	0.4603
	(-)	(-)	(-)	(-)	(-)
East	0.0866	0.0667	0.0616	0.0574	0.0476
	(-)	(-)	(-)	(-)	(-)
Midlands	0.1949	0.1621	0.1301	0.1013	0.1032
	(-)	(-)	(-)	(-)	(-)
South	0.2924	0.1621	0.2260	0.2095	0.2778
	(-)	(-)	(-)	(-)	(-)
Unemployment (months)	6.201	3.089	4.361	2.296	2.347
	(10.16)	(5.377)	(6.467)	(3.337)	(3.467)
$50 \leq$ Firm < 400	0.2996	0.2167	0.2466	0.1351	0.3095
	(-)	(-)	(-)	(-)	(-)
Firm ≥ 400	0.2167	0.2877	0.4426	0.2302	
	(-)	(-)	(-)	(-)	(-)
Change of Residence	0.0289	0.0485	0.0479	0.0979	0.0714
	(-)	(-)	(-)	(-)	(-)
Father's Occupation					
Non-Manual	0.1805	0.2636	0.3630	0.4831	0.6270
	(-)	(-)	(-)	(-)	(-)
Manual	0.5018	0.4621	0.4658	0.3682	0.2302
	(-)	(-)	(-)	(-)	(-)
$-\frac{\phi}{\Phi}$	1.3643	0.8486	1.8694	1.3032	1.5826
	(0.4643)	(0.8486)	(0.6502)	(0.3591)	(0.8152)
Observations	277	660	146	296	126

Standard deviations for continuous variables are in parentheses.

Table A.4.

Female Means by Occupational Category in Chapter Six.

Variable	Semi & Unskilled	Skilled	Other	Clerical	Prof & Man.
ln (Wage)	0.3528	0.3477	0.1345	0.4224	0.5976
	(0.2342)	(0.3932)	(0.4839)	(0.3280)	(0.3903)
Previous Experience	1.013	0.8652	1.142	3.049	3.292
	(1.702)	(1.283)	(1.810)	(1.875)	(1.934)
Experience	3.458	3.564	3.297	3.049	3.292
	(2.232)	(2.359)	(2.231)	(1.875)	(1.934)
Education	1.032	1.507	1.772	2.454	3.085
	(1.099)	(1.501)	(1.490)	(1.307)	(1.509)
Residence of Schooling					
Dublin City & Co.	0.3028	0.1507	0.3034	0.3826	0.3902
	(-)	(-)	(-)	(-)	(-)
East & Midlands	0.2191	0.1644	0.1517	0.1946	0.1829
	(-)	(-)	(-)	(-)	(-)
South	0.2869	0.5479	0.3448	0.2776	0.2805
	(-)	(-)	(-)	(-)	(-)
Current Residence					
Dublin City & Co.	0.3068	0.1507	0.3448	0.4137	0.5000
	(-)	(-)	(-)	(-)	(-)
East	0.0717	0.0959	0.0345	0.0726	0.0488
	(-)	(-)	(-)	(-)	(-)
Midlands	0.1513	0.0822	0.0965	0.1076	0.0854
	(-)	(-)	(-)	(-)	(-)
South	0.2829	0.5342	0.3448	0.2685	0.2561
	(-)	(-)	(-)	(-)	(-)
50 ≤ Firm < 400	0.5657	0.5342	0.1310	0.1984	0.3049
	(-)	(-)	(-)	(-)	(-)
Firm ≥400	0.2072	0.1507	0.2345	0.3048	0.4024
	(-)	(-)	(-)	(-)	(-)
Unemployment (months)	4.041	3.360	2.807	2.022	2.470
	(6.853)	(5.231)	(5.743)	(3.683)	(3.843)
Change of Residence	0.0239	0.0411	0.1034	0.0687	0.2927
	(-)	(-)	(-)	(-)	(-)
Father's Occupation					
Non-Manual	0.2390	0.2192	0.3172	0.3696	0.4643
	(-)	(-)	(-)	(-)	(-)
Manual	0.5857	0.5479	0.3586	0.3865	0.2195
	(-)	(-)	(-)	(-)	(-)
$-\frac{\phi}{\Phi}$	1.1996	2.1526	1.7629	0.5882	1.9911
	(0.4869)	(0.4952)	(0.3629)	(0.2536)	(1.0305)
Observations	251	73	145	771	82

Standard deviations for continuous variables are in parentheses.

Table A1.5

Means of Variables by Gender in Chapter Seven

Variable	Male	Female
Wage	0.4124	0.3845
Experience	3.441	3.196
Education	1.728	2.100
100 ≤ Firm < 500	0.1408	0.2015
Firm ≥500	0.2722	0.2762
Current Residence		
Dublin City & Co.	0.3254	0.3774
South	0.2842	0.2928
East & Mid.	0.2145	0.1796
Public	0.0936	0.1487
Observations	1506	1325

Bibliography

[1] Aigner, J. and Cain, G. (1977), Statistical Theories of Discrimination in Labour Markets, *Industrial and Labour Relations Review*, Vol. 30, pp. 175-187.

[2] Akerlof, G. (1982), Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics*, Vol. 97, pp. 543-569.

[3] Alchian, A. and Kessel, R. (1962), Competition, Monopoly and the Pursuit of Pecuniary Gains, in *Aspects of Labor Economics: A Conference of the Universities-National Bureau for Economic Research Committee*. Princeton, N.J.: Princeton University Press, pp. 156-175.

[4] Amemiya, T. (1981), Qualitative Response Models: A Survey. *Journal of Economic Literature*, Vol. 15, pp. 483-536.

[5] Amsden, A. (1980), The Economics of Women and Work, Introductory Chapter, *Penguin Books Ltd*.

[6] Arrow, K. (1972), The Theory of Discrimination, in *Discrimination in Labor Markets*, editors, Ashenfelter, O. and Rees, A.

[7] Ashenfelter, O. (1972), Racial Discrimination and Trades Unions, *Journal of Political Economy*, Vol. 80, pp. 435-464.

[8] _____________ and Hannan, T. (1986), Sex Discrimination and Product Market Competition: The Case of the Banking Industry, *Quarterly Journal of Economics*, Vol. 101, pp. 149-173.

[9] Becker, G. (1965), A Theory of the Allocation of Time, *Economic Journal*, Vol. 75, pp. 493-517.

[10] ___________ (1971), *The Economics of Discrimination*, University of Chicago Press, second edition.

[11] ___________ (1975), *Human Capital: a Theoretical and Empirical Analysis*, NBER, New York (second edition).

[12] __________ (1985), Human Capital, Effort, and the Sexual Division of Labour, *Journal of Labour Economics*, Vol. 3, (Papers and Proceedings) pp. s33-s58.

[13] Beller, A. (1985), Changes in the Sex Composition of US Occupations, 1960-1981, *Journal of Human Resources*, Vol. 20, pp. 235-250.

[14] Bergman, B. (1974), Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex, *Eastern Economic Journal*, Vol. 1, pp. 103-110.

[15] Blinder, A. (1973), Wage Discrimination: Reduced Form and Structural, *Journal of Human Resources*, Vol. 8, pp. 436-455.

[16] Boskin, M.J. (1974), A Conditional Logit Model of Occupational Choice. *Journal of Political Economy*, Vol. 82, pp. 389-398.

[17] Breusch, T.S. and Pagan, A.R. (1979), A Simple Test for Heteroscedasticity and Random Coefficient Variation, *Econometrica*, Vol. 47, pp. 1287-1294.

[18] Brown, R.S., Moon, M. and Zoloth, B.S. (1980), Incorporating Occupational Attainment in Studies of Male/Female Earnings Differentials, *Journal of Human Resources*, Vol. 15, pp. 3-28.

[19] Bulow, J. and Summers, L. (1986), A Theory of Dual Labour Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment, *Journal of Labor Economics*, Vol. 4, pp. 376-414.

[20] Cain, G. (1976), The Challenge of Segmented Labor Market Theories to Orthodox Theories, *Journal of Economic Literature*, Vol. 14, pp. 1215-1257.

[21] ________ (1986), The Economic Analysis of Labour Market Discrimination: A Survey, in *Handbook of Labour Economics*, Vol. 1, edited by O. Ashenfelter and R. Layard.

[22] Carnoy, M. and Rumberger, R. (1980), Segmentation in the US Labor Market: Its Effects on the Mobility and Earnings of Whites and Blacks, *Cambridge Journal of Economics*, Vol. 4, pp. 117-132.

[23] Chiplin, B. (1981), An Alternative Approach to the Measurement of Sex Discrimination: An Illustration from University Entrance. *Economic Journal* Vol. 91, pp. 988-997.

[24] ___________ and Sloane, P.J. (1976), Sex Discrimination in the Labour Market, (Macmillan).

[25] Chiswick, B. (1973), Racial Discrimination in the Labor Market: a Test of Alternative Hypotheses, *Journal of Political Economy*, Vol. 81, pp. 1330-1352.

[26] Chesher, A. and Irish, M. (1987), Residual Analysis in the Grouped and Censored Normal Linear Model, *Journal of Econometrics*, Vol. 34, pp. 33-61.

[27] Chowdhury, G. and Nickell, S. (1985), Hourly Earnings in the United States: Another Look at Unionisation, Schooling, Sickness, and Unemployment Using PSID Data, *Journal of Labor Economics*, Vol. 3, pp. 38-69.

[28] Comanor, W. (1973), Racial Discrimination in American Industry, *Economica*, Vol. 40, pp. 363-378.

[29] Corcoran, M. and Duncan, G. (1979), Work History, Labor Force Attachment and Earnings Differences Between Races and Sexes, *Journal of Human Resources*, Vol. 14, pp. 497-520.

[30] Dahlby, B. (1983), Adverse Selection and Statistical Discrimination, *Journal of Public Economics*, Vol. 20, pp. 121-130.

[31] Davidson, R. and MacKinnon, J. (1981), Several Tests for Model Specification in the Presence of Alternative Hypotheses, *Econometrica*, Vol. 49, pp.781-793.

[32] ________________________________ (1985), The Interpretation of Test Statistics, *Canadian Journal of Economics*, Vol. 18, pp.38-57.

[33] Dex, S. and Shaw, L. (1986), British and American Women at Work: Do Equal Opportunities Policies Matter? *Macmillan*.

[34] Dickens, W. and Lang, K. (1985), A Test of Dual Labor Market Theory, *American Economic Review*, Vol. 75, pp. 792-805.

[35] ______________________________ (1987), Where Have All the Good Jobs Gone ? Deindustrialisation and Labor Market Segmentation, in *Unemployment and the Structure of Labor Markets*, ed. Lang, K. and Leonard, S. (Basil Blackwell).

[36] ______________________________ (1988), The Reemergence of Segmented Labor Market Theory, *American Economic Review*, Vol. 73, (Papers and Proceedings) pp. 129-134.

[37] Doeringer, P. and Piore, M. (1971), Internal Labor Markets and Manpower Analysis, *Lexington*: Lexington Books.

[38] Dolton, P. (1984), Testing for Sex Discrimination in Maximum Likelihood Models, *Applied Economics* Vol. 16, pp. 225-235.

[39] _________ and Makepeace, G. (1986), Sample Selection and Male-Female Earnings Differentials in the Graduate Labour Market, *Oxford Economic Papers*, Vol. 38, pp. 317-341.

[40] ___________________________ (1987, a), Marital Status, Child Rearing and Earnings Differentials in the Graduate Labour Market, *Economic Journal*, Vol. 97, pp. 897-922.

[41] ___________________________ (1987, b), Interpreting Sample Selection Effects, *Economics Letters*, Vol. 24, pp. 373-379.

[42] ___________________________ and Van der Klaauw, W. (1987), Occupational Choice and Earnings Determination: The Role of Sample Selection and Non-Pecuniary Factors, *University of Bristol*, Discussion paper, # 87/186.

[43] Dubin, J. and M^c Fadden, D. (1984), An Econometric Analysis of Residential Electrical Appliance Holdings and Consumption, *Econometrica*, Vol. 52, pp.345-362.

[44] Duncan, O.D. and Duncan, B. (1955), A Methodological Analysis of Segregation Indices, *American Sociological Review*, Vol. 20, pp. 210-217.

[45] Duncan, G. (1983), Sample Selectivity as a Proxy Variable Problem: On the Use and Misuse of Gaussian Selectivity Corrections, in *New Approaches to Labour Unions* edited by Reid, J.

[46] _________ and Leigh, D. (1980), Wage Determination in the Union and Nonunion Sectors: A Sample Selectivity Approach, *Industrial and Labour Relations Review*, Vol. 34, pp.24-34.

[47] ______________________ (1985), The Endogeneity of Union Status: An Empirical Test, *Journal of Labor Economics*, Vol. 3, pp.385-402.

[48] Edgworth, F. (1922), Equal Pay to Men and Women, *Economic Journal*.

[49] Fawcett, M. (1918), Equal Pay for Equal Work, *Economic Journal*.

[50] Figart, D. (1987), Gender, Unions, and Internal Labor Markets: Evidence from the Public Sector in Two States. *American Economic Review*, Vol. 77, pp. 252-256.

[51] Fuchs, V. (1971), Differences in Hourly Earnings Between Men and Women, *Monthly Labor Review*, Vol. 94, pp. 9-15.

[52] Goldberger, A. (1982), Reverse Regression and Salary Discrimination, *Dept. of Economics, University of Wisconsin*.

[53] Goldfeld, S. and Quandt, R. (1976), Techniques for Estimating Switching Regressions, in *Studies in Non-Linear Estimation*, Cambridge, Ballinger.

[54] Greenhalgh, C. (1980), Male/Female Wage Differentials in Great Britain: Is Marriage an Equal Opportunity ?, *The Economic Journal*, Vol. 90, pp. 751-775.

[55] Gronau, R. (1977), Leisure, Home Production, and Work: the Theory of the Allocation of Time Revisited, *Journal of Political Economy*, Vol. 85, pp. 1099-1123.

[56] Hall, A. (1987), The Information Matrix Test for the Linear Model, *Review of Economic Studies*, Vol. 54, pp. 257-264.

[57] Haessel, W. and Palmer, J. (1978), Market Power and Employment Discrimination, *Journal of Human Resources*, Vol. 13, pp. 545-560.

[58] Hay, J. (1980), Occupational Choice and Occupational Earnings: Selectivity Bias in a Simultaneous Logit-OLS Model (unpublished Ph.D dissertation), *Yale University*

[59] Hausman, J. (1978), Specification Tests in Econometrics, *Econometrica*, Vol. 46, pp. 1251-1271.

[60] __________ and Taylor, W.E. (1981), Panel Data and Unobservable Individual Effects, *Econometrica*, Vol. 49, pp. 1377-1397.

[61] Heckman, J. (1976), The Common Structure of Statistical Models of Truncation, Self Selection and Limited Dependent Variables, and a Simple Estimator for Such Models, *The Annals of Economic and Social Measurement*, Vol. 5, pp. 475-492.

[62] __________ (1979), Sample Selection Bias as a Specification Error, *Econometrica*, Vol. 47, pp. 153-161.

[63] __________ and Hotz, V.J. (1986), An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality, *Journal of Human Resources*, Vol. 21, pp. 507-542.

[64] Hill, M.A. (1983), Female Labor Force Participation in Developing and Developed Countries-Consideration of the Informal Sector, *Review of Economics and Statistics*, Vol. 45, pp. 459-468.

[65] Johnson, G. and Solon, G. (1984), Pay Differences Between Women's and Men's Jobs: the Empirical Foundation of Comparable Worth Legislation, *NBER*.

[66] Jones, S. (1987), Minimum Wage Legislation in a Dual Labor Market, *European Economic Review*, Vol. 31, pp. 1229-1246.

[67] Kahn, L. and Sherer, P. (1988), Racial Differences in Professional Basketball Players' Compensation, *Journal of Labour Economics*, Vol. 6, pp. 40-61.

[68] Kamalich, R. and Polachek, S. (1982), Discrimination: Fact or Fiction ? An Examination Using an Alternative Approach, *Southern Economic Journal*, Vol. 49, pp. 450-461.

[69] Katz, L. (1986), Efficiency Wage Theories: A Partial Evaluation, *NBER Macroeconomics Annual*, Cambridge: MIT Press, pp. 235-276.

[70] Kiefer, N.M. and Salmon, M. (1983), Testing Normality in Econometric Models, *Economics Letters*, Vol. 11, pp. 123-127.

[71] Killingsworth, M. (1983), Labor Supply, *Cambridge University Press*.

[72] Lancaster, K. (1966), A New Approach to Consumer Theory, *Journal of Political Economy*, Vol. 74, pp. 132-157.

[73] Lee, L-F. (1978), Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables, *International Economic Review*, Vol. 19, pp. 415-433.

[74] _________ (1982), Some Approaches to the Correction of Selectivity Bias, *Review of Economic Studies*, Vol. 49, pp. 355-372.

[75] _________ (1983), Generalised Econometric Models with Selectivity, *Econometrica*, Vol. 51, pp. 507-512.

[76] _________ Maddala, G. and Trost, R. (1980), Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity, *Econometrica*, Vol. 48, pp. 491-503.

[77] Leigh, D. (1976), Occupational Advancement in the Late Sixties, An Indirect Test of the Dual labour Market Hypothesis, *Journal of Human Resources*, Vol. 11, pp. 155-171.

[78] Lewis, H. (1986), Union Relative Wage Effects: A Survey *University of Chicago Press.*

[79] Lucifora, C. and Reilly, B. (1988), Wages and Sex Discrimination in Italy: an Empirical Analysis Using Micro Data, *mimeo* Universita Cattolica di Milano.

[80] Luksetich, W. (1979), Market Power and Sex Discrimination in White-Collar Employment, *Review of Social Economy*, Vol. 37, pp. 211-224.

[81] Lundberg, S. and Startz, R. (1983), Private Discrimination and Social Intervention in Competitive Labour Markets, *American Economic Review*, Vol. 73, pp. 340-347.

[82] Maddala, G. (1983), Limited-Dependent and Qualitative Variables in Econometrics, *Cambridge University Press.*

[83] ___________ (1986), Limited-Dependent and Qualitative Variables in Econometrics, *Cambridge University Press.*

[84] Madden, J. (1987), Gender Differences in the Cost of Displacement: An Empirical Test of Discrimination in the Labor Market, *American Economic Review*, Vol. 77, pp. 246-251.

[85] Malcolmson, J. (1984), Work Incentives, Hierarchy, and Internal Labor Markets, *Journal of Political Economy*, Vol. 92, pp. 486-507.

[86] M^c Donald, I. and Solow, R. (1981), Wage Bargaining and Employment, *American Economic Review*, Vol. 71, pp. 896-908.

[87] ______________________________(1985), Wages and Employment in a Segmented Labour Market, *Quarterly Journal of Economics*, Vol. 93, pp. 312-330.

[88] M^c Fadden, D. (1973), Conditional Logit Analysis of Qualitative Choice Behaviour, in *Frontiers in Econometrics*, ed. Zarembka, P. (New York).

[89] Melino, A. (1982), Testing for Sample Selection Bias, *Review of Economic Studies*, Vol. 49, pp. 151-153.

[90] Miller, P.W. (1987), The Wage Effect of the Occupational Segregation of Women in Britain, *The Economic Journal*, Vol. 97, pp. 885-896.

[91]__________ and Volker, P.A. (1985), On the Determination of Occupational Attainment and Mobility, *Journal of Human Resources*, Vol. 20, pp. 197-213.

[92] Mincer, J. (1974) , Schooling, Experience and Earnings, *National Bureau of Economic Research*.

[93] __________ (1985), Intercountry Comparisons of Labor Force Trends and of Related Developments: An Overview, *Journal of Labour Economics*, Vol. 3, (Papers and Proceedings) pp. s1-s32.

[94] __________ and Polachek, S. (1974), Family Investments in Human Capital: Earnings of Women, *Journal of Political Economy*, Vol. 82, pp. S76-S108.

[95] Oaxaca, R. (1973), Male/Female Wage Differentials in Urban Labour Markets, *International Economic Review*, Vol. 14, pp. 693-709.

[96] Olsen, R. (1980), A Least Squares Correction for Selectivity Bias, *Econometrica*, Vol. 48, pp. 1815-1820.

[97] __________ (1982), Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator, *International Economic Review*, Vol. 23, pp. 223-240.

[98] Oster, S. (1975), Industry Differences in Discrimination Against Women, *Quarterly Journal of Economics*, Vol. 89, pp. 215-229.

[99] Osterman, P. (1975), An Empirical Study of Labor Market Segmentation, *Industrial and Labor Relations Review*, Vol. 28, pp. 508-523.

[100] Pagan, A. (1984), Model Evaluation by Variable Addition, Chapter 3 in *Econometrics and Quantitative Economics*, editors, Hendry, D. and Wallis, K.F., (Blackwell).

[101] Papanicolau, J., and Psacharopoulos, G., (1976), Socioeconomic Background, Schooling and Monetary Rewards in the United Kingdom, *Economica*, Vol. 46, pp.435-439.

[102] Phelps, E. (1972), The Statistical Theory of Racism and Sexism, *American Economic Review*, Vol. 62, pp. 659-661.

[103] Piore, M. (1974), Primary and Secondary Labour Markets: A Critique of the Dual Approach, *Brookings Papers on Economic Activity*, Vol. 3, pp. 684-688.

[104] ________ (1980), Dualism as a Response to Flux and Uncertainty, in *Dualism and Discontinuity in Industrial Societies*, ed. Berger, S. and Piore, M. New York: *Cambridge University Press*.

[105] ________ (1983), Labor Market Segmentation: To What Paradigm Does it Belong ?, *American Economic Review*, Vol. 73, (Papers and Proceedings) pp. 249-253.

[106] Polachek, S. (1975), Potential Biases in Measuring Male-Female Discrimination, *Journal of Human Resources*, Vol. 11, pp. 203-229.

[107] ________ (1981), Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure, *Review of Economics and Statistics*, Vol. 63, pp. 60-69.

[108] Ramsey, J. (1969), Tests for Specification in Classical Linear Least Squares Regression analysis, *Journal of Royal Statistical Society*, Vol. 31, pp. 350-371.

[109] ________ and Schmidt, P. (1976), Some Further Results on the Use of OLS and BLUS Residuals in Specification Error, *Journal of the American Statistical Association*, Vol. 67, pp. 389-390.

[110] Reich, M., Gordon, D. and Edwards, R. (1973), A Theory of Labour Market Segmentation, *American Economic Review*, Vol. 63, pp. 359-365.

[111] Reilly, B. (1987), Wages, Sex Discrimination and the Irish Labour Market for Young Workers, *Economic and Social Review*, Vol. 18, pp. 271-305.

[112] Reimers, C.W. (1983), Labour Market Discrimination Against Hispanic and Black Men, *Review of Economics and Statistics*, Vol. 45, pp. 570-579.

[113] Robinson, C. and Tomes, N. (1982), Self-Selection and Interprovincial Migration in Canada, *Canadian Journal of Economics*, Vol. 15, pp. 474-502.

[114] Roemer, J. (1979), Divide and Conquer, Microfoundation of a Marxian Theory of Wage Discrimination, *Bell Journal of Economics*, Vol. 10, pp. 695-705.

[115] Rosen, S. (1974), Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy*, Vol. 82, pp. 34-55.

[116] Schiller, B. (1977), Relative earnings Mobility in the United States, *American Economic Review*, Vol. 67 pp. 926-941.

[117] Schmidt, P. and Strauss, R.P. (1976), The Predictions of Occupation Using Multiple Logit Models, *International Economic Review*, Vol. 16, pp. 471-486.

[118] Sexton, J., Whelan, B. and Dillon, M. (1982), A Survey of Youth Employment and the Transition from Education to Working Life, *ESRI, Dublin*.

[119] Shapiro, C. and Stiglitz, J. (1984), Involuntary Unemployment as a Worker Discipline Device, *American Economic Review*, Vol. 74, pp. 433-444.

[120] Shephard W. and Levin, S. (1973), Managerial Discrimination in Large Firms, *Review of Economics and Statistics*, Vol. 55, pp. 412-422

[121] Siebert, W. and Sloane, P. (1981), The Measurement of Sex and Marital Status Discrimination at the Workplace, *Economica*, Vol. 48, pp. 125-141.

[122] Sloane, P. (1985), Discrimination in the Labour Market, Chapter 2 in *Labour Economics*, by Carline, D., Pissarides, C.A., Siebert, W.S. and Sloane P.J. (Longman).

[123] Solon, G. (1983). Errors in Variables and Reverse Regression in the Measurement of Wage Discrimination, *mimeo*, Princeton University.

[124] Spence, M. (1973), Job Market Signaling, *Quarterly Journal of Economics*, Vol. 87, pp. 355-374.

[125] Stewart, M. (1983), Relative Earnings and Individual Union Membership in the United Kingdom, *Economica*, Vol. 50, pp. 111-125.

[126] __________ (1987), Collective Bargaining Arrangements, Closed Shops and Relative Pay, *The Economic Journal*, Vol. 97, pp.140-156.

[127] __________ (1988), Union Wage Differentials, Product Market Influences and the Division of Rents, *Mimeo*, University of Warwick.

[128] __________ and Greenhalgh, C. (1984), Work History Patterns and Occupational Attainment of Women, *Economic Journal*, Vol. 94, pp. 493-519.

[129] ____________ and Wallis, K.F. (1981), *Introductory Econometrics*, Blackwell, second edition.

[130] Stiglitz, J. (1974), Theories of Discrimination and Economic Policy, in *Patterns of Racial Discrimination*, ed. Von Furstenberg, G., Horowitz, A. and Harrison, B. (Lexington).

[131] Taubman, P. (1976), Earnings, Education, Genetics and Environment, *The Journal of Human Resources*, Vol. 11, pp.447-461.

[132] Thurow, L. (1968), The Determinants of the Occupational Distribution of Negroes in *The Education and Training of Racial Minorities: Proceedings of a Conference*. Madison: Center for Studies in Vocational Education, University of Wisconsin.

[133] ____________ (1975), Generating Inequality, *Basic Books*, New York.

[134] ____________ and Lucas, R. (1972), The American Distribution of Income. Joint Economic Committee Print, 92d Congress, Second Session.

[135] Trost, R. and Lee, L-F (1983), Technical Earnings and Training: A Polychotomous Choice Model with Selectivity, *Review of Economics and Statistics*, Vol. 45, pp. 151-156.

[136] Williamson, O., Wachter, L,. and Harris, J., (1975), Understanding the Employment Relationship: the Analysis of Idiosyncratic Exchange, *Bell Journal of Economics*, Vol. 3, pp. 250-278.

[137] Wachter, M. (1974), Primary and Secondary Labor Markets: A Critique of the Dual Approach, *Brookings Papers on Economic Activity*, Vol. 3, pp. 637-680.

[138] Walsh, B.M. and Whelan, B.J. (1976), A Micro-Economic Study of Earnings in Ireland, *Economic and Social Review*, Vol. 7, pp. 199-217.

[139] Webb, B. (1918), The Wages of Men and Women: Should They be Equal ? *George Allen & Unwin*.

[140] White, H. (1980), A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica*, Vol. 48, pp. 817-838.

[141] ____________ (1982, a), Maximum Likelihood Estimation of Misspecified Models, *Econometrica*, Vol. 50, pp. 1-25.

[142] ________ (1982, b), Instrumental Variables Regression with Independent Observations, *Econometrica*, Vol. 50, pp. 483-496.

[143] Yule, G. and Kendall, M. (1950), An Introduction to the Theory of Statistics, *Griffin & Co.* London, (14th edition.

[144] Zabalza, A. (1983), The CES Utility Function, Non-Linear Budget Constraints and Labour Supply. Results on Female Participation and Hours, *Economic Journal*, Vol. 93, pp. 312-330.

[145] ________ and Tzannatos, Z. (1985), The Effect of Anti-Discriminatory Legislation on Relative Pay and Employment, Ch. 4 in Women and Equal Pay, edited by Zabalza, A. and Tzannatos, Z. *Cambridge University Press*

[146] ________ and Arrufat, J.L. (1985), The Extent of Sex Discrimination in Great Britain, Ch. 5 in Women and Equal Pay, edited by Zabalza, A. and Tzannatos, Z. *Cambridge University Press*

[147] Zellner, H. (1972), Discrimination Against Women, Occupational Segregation and the Relative Wage, *American Economic Review*, Vol. 62, pp. 157-160.

[148] Zucker, L. and Rosenstein, C. (1981), Taxonomies of Institutional Structure: Dual Economy Reconsidered, *American Sociological Review*, Vol. 46, pp. 869-884.

TITLE

Gender Wage Differentials and the Labour Market for Young Workers:

An Empirical Analysis Using Data for Ireland

AUTHOR

Barry Reilly

INSTITUTION and DATE

University of Warwick

March 1989